

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

PCT/JP01/01982

13.03.01

日本国特許庁

PATENT OFFICE
JAPANESE GOVERNMENT

REC'D 26 MAR 2001

WIPO

PCT

10/009119

JP01/1982

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

2000年 3月13日

出願番号

Application Number:

特願2000-068720

出願人

Applicant(s):

ソニー株式会社

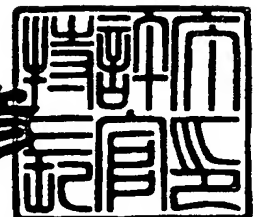
PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1 (a) OR (b)

2000年 6月 2日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



出証番号 出証特2000-3040322

【書類名】 特許願
【整理番号】 0000191017
【提出日】 平成12年 3月13日
【あて先】 特許庁長官殿
【国際特許分類】 G06F 19/00

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社
内

【氏名】 ピーター クーン

【特許出願人】

【識別番号】 000002185
【氏名又は名称】 ソニー株式会社
【代表者】 出井 伸之

【代理人】

【識別番号】 100082131
【弁理士】
【氏名又は名称】 稲本 義雄
【電話番号】 03-3369-6479

【手数料の表示】

【予納台帳番号】 032089
【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【包括委任状番号】 9708842

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ビデオ／オーディオシグナル処理方法、およびビデオ／オーディオシグナル処理装置

【特許請求の範囲】

【請求項1】 ビデオ／オーディオシグナルを供給し、処理するビデオ／オーディオシグナル処理方法であって、

コード変換目標ビットストリームパラメータを記述するステップと、
コード変換ヒントメタデータを抽出するステップと、
コード変換ヒントメタデータを記憶するステップと、
セグメントへの視聴覚素材を分離するステップと、
前記視聴覚セグメントに対してコード変換ヒントメタデータを関連させるステップと、

コード変換するステップと

を含むことを特徴とするビデオ／オーディオシグナル処理方法。

【請求項2】 請求項1記載のコード変換目標ビットストリームパラメータ記述ステップは、

圧縮されたイメージの2番目のビットストリームのビットレートを定義するステップと、

圧縮されたイメージの2番目のビットストリームのピクチャのサイズの定義するステップと、

圧縮されたイメージの2番目のビットストリームの2番目あたりフレームの数を定義するステップと、

圧縮されたイメージの2番目のビットストリームのピクセルの縦横比を定義するステップと、

圧縮されたイメージの2番目のビットストリームのピクセルのうちのそれぞれの色深さを定義するステップと、

もし圧縮されたイメージの2番目のビットストリームのための進歩的なフォーマットの使用定義するステップと、

もし圧縮されたイメージの2番目のビットストリームのための織り混ぜられた

フォーマットの使用定義するステップと、

もしフレームが描くならば定義することは、圧縮されたイメージの2番目のビットストリームのために使うステップと、

もしフィールドが描くならば定義することは、圧縮されたイメージの2番目のビットストリームのために使うステップと、

圧縮されたイメージの2番目のビットストリームの圧縮方法を定義するステップと

を含むことを特徴とする請求項1記載のビデオ／オーディオシグナル処理方法

。 【請求項3】 前記コード変換目標記述ステップは、MPEG（映画専門家グループ）により定義される圧縮標準を使用するステップをさらに含む

ことを特徴とする請求項2記載のビデオ／オーディオシグナル処理方法。

【請求項4】 前記コード変換目標記述ステップは、ITU-T（国際的なテレコミュニケーション連合技術標準グループ）により定義される圧縮標準を使用するステップをさらに含む

ことを特徴とする請求項2記載のビデオ／オーディオシグナル処理方法。

【請求項5】 前記のコード変換ヒント抽出ステップは、最初のGOP構造を持っている圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの獲得最初の動作情報するステップと、

最初のセグメンテーションの獲得テクスチャー／エッジ情報するステップと、

こぶしビットストリームからの獲得機能ポイントおよび関連した動作情報するステップと、

興味情報の領域を最初のビットストリームから得るステップと

を含むことを特徴とする請求項1記載のビデオ／オーディオシグナル処理方法

。 【請求項6】 前記コード変換ヒント抽出ステップは、

最初のGOP構造を持っている圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの獲得最初の動作情報するステップと、
コード変換ヒントメタデータとしてこの動作情報を蓄えるステップと
を含むことを特徴とする請求項 5 記載のビデオ／オーディオシグナル処理方法

【請求項 7】 前記コード変換ヒント抽出ステップは、パラメータの動作モデルのパラメータとしての象徴動作関連コード変換ヒントメタデータするステップ

を含むことを特徴とする請求項 5 記載のビデオ／オーディオシグナル処理方法

【請求項 8】 前記コード変換ヒント抽出ステップは、*furhter* 次の矩形ビデオフレーム内のグローバルな動作を説明しているパラメータの動作モデルを雇用するステップ

を含むことを特徴とする請求項 7 記載のビデオ／オーディオシグナル処理方法

【請求項 9】 前記コード変換ヒント抽出ステップは、任意の形の定義された領域内の動作を説明しているパラメータの動作モデルを雇用するステップ

を含むことを特徴とする請求項 7 記載のビデオ／オーディオシグナル処理方法

【請求項 1 0】 前記コード変換ヒント抽出ステップは、MPEG-4 内で使ったように、任意の形の定義された領域内の動作を説明しているパラメータの動作モデルを雇用するステップ

を含むことを特徴とする請求項 9 記載のビデオ／オーディオシグナル処理方法

【請求項 1 1】 前記コード変換ヒント抽出ステップは、動きベクトルの配列としての象徴動作関連コード変換ヒントメタデータは、圧縮されたイメージデータの最初のビットストリームに含むステップ

を含むことを特徴とする請求項 5 記載のビデオ／オーディオシグナル処理方法

【請求項 1 2】 前記コード変換ヒント抽出ステップは、動きベクトルの配

列としての象徴動作関連コード変換ヒントメタデータは、圧縮されたイメージデータの最初のビットストリームに含まれている動きベクトルから派生するステップ

を含むことを特徴とする請求項 5 記載のビデオ／オーディオシグナル処理方法

【請求項 1 3】 前記コード変換ヒント抽出ステップは、機能のリストとしての象徴動作関連コード変換ヒントメタデータは、

関連した動きベクトルによって示し、それは次のフレーム内で追跡されるステップ

を含むことを特徴とする請求項 5 記載のビデオ／オーディオシグナル処理方法

【請求項 1 4】 前記コード変換ヒント抽出ステップは、機能のリストとしての象徴動作関連コード変換ヒントメタデータは、関連した動きベクトルによって示し、それは、次のフレーム内の恣意的に具体化された領域内に追跡されるステップ

を含むことを特徴とする請求項 5 記載のビデオ／オーディオシグナル処理方法

【請求項 1 5】 前記コード変換ヒント抽出ステップは、DCT-係数または手段（平均、最小、最高点、相違、および標準偏差のような）のリストとしてのテクスチャ関連コード変換ヒントメタデータの表現は、そこから派生するステップ

を含むことを特徴とする請求項 5 記載のビデオ／オーディオシグナル処理方法。

【請求項 1 6】 前記コード変換ヒント抽出ステップは、そのリストとしてのエッジ関連コード変換ヒントメタデータの表現、またはそこから手段（平均、最小、最高点、相違、および標準偏差のような）が派生するステップ

を含むことを特徴とする請求項 5 記載のビデオ／オーディオシグナル処理方法

【請求項 1 7】 前記コード変換ヒント抽出ステップは、機能ポイントおよ

び関連した動作関連コード変換ヒントメタデータがリストであるのを見せるステップ

を含むことを特徴とする請求項5記載のビデオ／オーディオシグナル処理方法

【請求項18】 前記コード変換ヒント抽出ステップは、メートル法である複雑さとしての符号化複雑さ関連コード変換ヒントメタデータの表現は、次のフレーム内で追跡されたfeaturepointsのライフ時間リストから派生するステップと、 前記1つのフレームから次のフレームに、失われるおよび新しいfeaturepointsの数を使うことによってするステップと

を含むことを特徴とする請求項5記載のビデオ／オーディオシグナル処理方法

【請求項19】 前記コード変換ヒントメタデータ記憶ステップは、いくつかの状況のためのコード変換ヒントメタデータを含んでいるバッファを運営するステップ

を含むことを特徴とする請求項1記載のビデオ／オーディオシグナル処理方法

【請求項20】 前記コード変換ヒントメタデータを記憶するステップは、いくつかの目標機器のために、個々一般的コード変換ヒントメタデータを蓄えるステップ

を含むことを特徴とする請求項19記載のビデオ／オーディオシグナル処理方法。

【請求項21】 前記コード変換ヒントメタデータを記憶するステップは、変化場面活動の視聴覚セグメントのために、一般コード変換ヒントメタデータを蓄えるステップ

を含むことを特徴とする請求項19記載のビデオ／オーディオシグナル処理方法。

【請求項22】 前記セグメントへの視聴覚素材を分離するステップは、機能を使うことは、関連した動きベクトルによって示すステップと、機能を追跡することを示し、機能ポイントのライフ時間を守るステップと、

前記 1 つのフレームから次のフレームに追跡できなかった機能ポイントの数のコード変換扶養家族のための新しい視聴覚セグメントを決定するステップとを含むことを特徴とする請求項 1 記載のビデオ／オーディオシグナル処理方法。

【請求項 2 3】 前記視聴覚セグメントに対してコード変換ヒントメタデータを関連させるステップは、

クレーム毎の新しい機能のポイントの数を計算するステップと、
新しい機能のポイントの数が所定の閾値を越えるか否かを決定するステップと

いくつかのコード変換ヒントステートのうちのこの計測に基づき選択するステップと

を含むことを特徴とする請求項 1 記載のビデオ／オーディオシグナル処理方法。

【請求項 2 4】 前記コード変換ヒント抽出ステップは、

最初の GOP 構造を持っている圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの抽出コード変換ヒントメタデータするステップと、

コード変換を容易にするように最初のビットストリームに関連したコード変換ヒントメタデータを利用するステップと、

2 番目のビットストリームのアウトプットを提供するステップと

を含むことを特徴とする請求項 1 記載のビデオ／オーディオシグナル処理方法。

【請求項 2 5】 前記コード変換ヒント抽出ステップは、前記ビットストリームの時間的なセグメントに関連したコード変換ヒントメタデータを利用するステップ

を含むことを特徴とする請求項 2 4 記載のビデオ／オーディオシグナル処理方法。

【請求項 2 6】 前記コード変換ヒント抽出ステップは、前記ビットストリ

ームの空間のセグメントに関連したコード変換ヒントメタデータを利用するステップ

を含むことを特徴とする請求項 24 記載のビデオ／オーディオシグナル処理方法。

【請求項 27】 前記コード変換ヒント抽出ステップは、前記 GOP 構造と違う 2 番目の GOP 構造を持っている圧縮されたイメージデータの 2 番目のビットストリームのための 2 番目の動作情報を補外するために、コード変換ヒントメタデータに含まれている動作情報を利用するステップと、

前記 2 番目のビットストリームのアウトプットを提供するステップと
を含むことを特徴とする請求項 24 記載のビデオ／オーディオシグナル処理方法。

【請求項 28】 前記コード変換ヒント抽出ステップは、最初のビットストリームのビットレートが 2 番目のビットストリームのビットレートと違うことで 2 番目のビットストリームのビットレートをコントロールしているステップと、

前記ビットストリーム従って最初のビットストリームに表されているピクチャのサイズと違う 2 番目のビットストリーム展示品サイズに表されているピクチャに表されているピクチャのサイズを調整しているステップにおいて要求されるステップ

を含むことを特徴とする請求項 24 記載のビデオ／オーディオシグナル処理方法。

【請求項 29】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームに表されている写真が最初のビットストリームに表されている写真のサイズと違うサイズを表すステップと、

最初のビットストリームに表されている写真のサイズを調整しているステップと

を含むことを特徴とする請求項 28 記載のビデオ／オーディオシグナル処理方法。

【請求項 30】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームが、最初のビットストリームに表されている写真がフレーム写真として

符号化される時のフィールド写真を見せる写真を符号化しているステップ

を含むことを特徴とする請求項 2 9 記載のビデオ／オーディオシグナル処理方法。

【請求項 3 1】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームが、最初のビットストリームに表されている写真がフィールド写真として符号化される時のフレーム写真を見せる写真を符号化しているステップ

を含むことを特徴とする請求項 2 9 記載のビデオ／オーディオシグナル処理方法。

【請求項 3 2】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームに表されている写真が、織り混ぜられた連鎖として出力されるように、最初のビットストリームに表されている写真が進歩的な連鎖と認められる時の最初のビットストリームによって写真を織り混ぜているステップ

を含むことを特徴とする請求項 2 9 記載のビデオ／オーディオシグナル処理方法。

【請求項 3 3】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームに表されている写真が、織り混ぜられた連鎖として出力されるように、最初のビットストリームに表されている写真が進歩的な連鎖と認められる時の最初のビットストリームによって写真を不織り混ぜているステップ

を含むことを特徴とする請求項 2 9 記載のビデオ／オーディオシグナル処理方法。

【請求項 3 4】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームが、最初のビットストリームに表されている写真がフレーム写真として符号化される時のフィールド写真を見せる符号化写真のステップ

を含むことを特徴とする請求項 2 4 記載のビデオ／オーディオシグナル処理方法。

【請求項 3 5】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームが、最初のビットストリームに表されている写真がフィールド写真として符号化される時のフレーム写真を見せる符号化写真のステップ

を含むことを特徴とする請求項 2 4 記載のビデオ／オーディオシグナル処理方

法。

【請求項 3 6】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームに表されている写真が、織り混ぜられた連鎖として出力されるように、最初のビットストリームに表されている写真が進歩的な連鎖と認められる時の最初のビットストリームに表されている組み合わせ写真のステップ

を含むことを特徴とする請求項 2 4 記載のビデオ／オーディオシグナル処理方法。

【請求項 3 7】 前記コード変換ヒント抽出ステップは、2 番目のビットストリームに表されている写真が、進歩的な連鎖として出力されるように、最初のビットストリームに表されている写真が織り混ぜられた連鎖と認められる時の最初のビットストリームに表されている不組み合わせ写真のステップ

を含むことを特徴とする請求項 2 4 記載のビデオ／オーディオシグナル処理方法。

【請求項 3 8】 最初のサイズの写真を表している圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの動作関連転送コード化ヒント隠喩データの抽出するステップと、

動作関連転送コード化ヒント隠喩データの格納するステップと、

最初のサイズと違う 2 番目のサイズの写真を表している圧縮されたイメージデータの 2 番目のビットストリームのための 2 番目の動作情報を補外するために、

格納された最初の動作関連転送コード化ヒント隠喩データを利用するステップと、

さらに 2 番目のビットストリームのアウトプットを提供するステップと

を含むことを特徴とするコード変換方法。

【請求項 3 9】 説明している圧縮されたイメージデータの受領最初のビットストリームは、織り混ぜられた連鎖を定義するのを描くステップと、

最初のビットストリームからの動作関連コード変換ヒントメタデータの抽出するステップと、

動作関連コード変換ヒントメタデータの格納するステップと、

説明している圧縮されたイメージデータの2番目のビットストリームのための2番目の動作情報を補外するために、格納された最初の動作関連コード変換ヒントメタデータを利用することにより、進歩的な連鎖を定義するのが描くステップと、

さらに2番目のビットストリームのアウトプットを提供するステップとを含むことを特徴とするコード変換方法。

【請求項40】 説明している圧縮されたイメージデータの中で最初のビットストリームを受け取ることにより、進歩的な連鎖を定義するのが描くステップと、

最初のビットストリームからの動作関連コード変換ヒントメタデータの抽出するステップと、

動作関連コード変換ヒントメタデータの格納するステップと、

説明している圧縮されたイメージデータの2番目のビットストリームのための2番目の動作情報を補外するために、格納された最初の動作関連コード変換ヒントメタデータを利用することにより、織り混ぜられた連鎖を定義するのが描くステップと、

前記2番目のビットストリームのアウトプットを提供するステップとを含むことを特徴とするコード変換方法。

【請求項41】 フレームピクチャを表している圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの動作関連コード変換ヒントメタデータの抽出するステップと、

動作関連コード変換ヒントメタデータの格納するステップと、

フィールドピクチャを表している圧縮されたイメージデータの2番目のビットストリームのための2番目の動作情報を補外するために、格納された最初の動作関連コード変換ヒントメタデータを利用するステップと、

前記2番目のビットストリームのアウトプットを提供するステップとを含むことを特徴とするコード変換方法。

【請求項42】 フィールドピクチャを表している圧縮されたイメージデー

タの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの動作関連コード変換ヒントメタデータの抽出するステップと、

動作関連コード変換ヒントメタデータの格納するステップと、

フレームピクチャを表している圧縮されたイメージデータの2番目のビットストリームのための2番目の動作情報を補外するために、格納された最初の動作関連コード変換ヒントメタデータを利用するステップと、

前記2番目のビットストリームのアウトプットを提供するステップと

を含むことを特徴とするコード変換方法。

【請求項43】 主要なイメージを表している圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの動作関連コード変換ヒントメタデータの抽出するステップと、

動作関連コード変換ヒントメタデータの格納するステップと、

主要なイメージの部分を表している圧縮されたイメージデータの2番目のビットストリームのための2番目の動作情報を補外するために、格納された最初の動作関連コード変換ヒントメタデータを利用するステップと、

前記2番目のビットストリームのアウトプットを提供するステップと

を含むことを特徴とするコード変換方法。

【請求項44】 GOP構造、ピクチャサイズ、ビットレート、およびフレーム、またはフィールドピクチャフォーマット、および進歩的なまたは織り混ぜられた連鎖を含むコード化パラメータの大多数を持っている圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの動作関連コード変換ヒントメタデータの抽出するステップと、

動作関連コード変換ヒントメタデータの格納するステップと、

コード化パラメータそのようなもの2番目のビットストリームのコード化パラメータのうちの1つ以上が最初のビットストリームのコード化パラメータと違っている圧縮されたイメージデータの2番目のビットストリームの

ための 2 番目の動作情報を補外するために、格納された最初の動作関連コード変換ヒントメタデータを利用するステップと、

前記 2 番目のビットストリームのアウトプットを提供するステップと
を含むことを特徴とするコード変換方法。

【請求項 4 5】 最初のサイズのピクチャを表している圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの機能ポイント動作関連コード変換ヒントメタデータを抽出するステップと、

機能ポイント動作関連コード変換ヒントメタデータの格納するステップと、

最初のサイズと違う 2 番目のサイズのピクチャを表している圧縮されたイメージデータの 2 番目のビットストリームのための 2 番目の動作情報を補外するために、格納された最初の機能ポイント動作関連コード変換ヒントメタデータを利用するステップと、

前記 2 番目のビットストリームのアウトプットを提供するステップと
を含むことを特徴とするコード変換方法。

【請求項 4 6】 説明している圧縮されたイメージデータの受領最初のビットストリームは、織り混ぜられた連鎖を定義するのを描くステップと、

最初のビットストリームからの機能ポイント動作関連コード変換ヒントメタデータの抽出するステップと、

機能ポイント動作関連コード変換ヒントメタデータの格納するステップと、

説明している圧縮されたイメージデータの 2 番目のビットストリームのための 2 番目の動作情報を補外するために、格納された最初の機能ポイント動作関連コード変換ヒントメタデータを利用することにより、進歩的な連鎖を定義するのが描くステップと、

前記 2 番目のビットストリームのアウトプットを提供するステップと
を含むことを特徴とするコード変換方法。

【請求項 4 7】 説明している圧縮されたイメージデータの中で最初のビットストリームを受け取ることにより、進歩的な連鎖を定義するのが描くステップと、

最初のビットストリームからの機能ポイント動作関連コード変換ヒントメタデータの抽出するステップと、

機能ポイント動作関連コード変換ヒントメタデータの格納するステップと、
説明している圧縮されたイメージデータの2番目のビットストリームのための2番目の動作情報を補外するために、格納された最初の機能ポイント動作関連コード

~~変換ヒントメタデータを利用することにより、織り混ぜられた連鎖を定義する~~
のが描くステップと、

前記2番目のビットストリームのアウトプットを提供するステップと
を含むことを特徴とするコード変換方法。

【請求項48】 フレームピクチャを表している圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの機能ポイント動作関連コード変換ヒントメタデータの抽出するステップと、

機能ポイント動作関連コード変換ヒントメタデータの格納するステップと、

フィールドピクチャを表している圧縮されたイメージデータの2番目のビットストリームのための2番目の動作情報を補外するために、格納された最初の機能ポイント動作関連コード変換ヒントメタデータを利用するステップと、

前記2番目のビットストリームのアウトプットを提供するステップと
を含むことを特徴とするコード変換方法。

【請求項49】 フィールドピクチャを表している圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの機能ポイント動作関連コード変換ヒントメタデータの抽出するステップと、

機能ポイント動作関連コード変換ヒントメタデータの格納するステップと、

フレームピクチャを表している圧縮されたイメージデータの2番目のビットストリームのための2番目の動作情報を補外するために、格納された最初の動作関連コード変換ヒントメタデータを利用するステップと、

前記2番目のビットストリームのアウトプットを提供するステップと
を含むことを特徴とするコード変換方法。

【請求項 5 0】 主要なイメージを表している圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの機能ポイント動作関連コード変換ヒントメタデータの抽出するステップと、

機能ポイント動作関連コード変換ヒントメタデータの格納するステップと、

主要なイメージの部分を表している圧縮されたイメージデータの 2 番目のビットストリームのための 2 番目の動作情報を補外するために、格納された最初の機能ポイント動作関連コード変換ヒントメタデータを利用するステップと、

前記 2 番目のビットストリームのアウトプットを提供するステップと

を含むことを特徴とするコード変換方法。

【請求項 5 1】 GOP 構造、ピクチャサイズ、ビットレート、およびフレーム、またはフィールドピクチャフォーマット、および進歩的なまたは織り混ぜられた連鎖を含むコード化パラメータの大多数を持っている圧縮されたイメージデータの中で最初のビットストリームを受け取るステップと、

最初のビットストリームからの機能ポイント動作関連コード変換ヒントメタデータの抽出するステップと、

機能ポイント動作関連コード変換ヒントメタデータの格納するステップと、

コード化パラメータそのようなもの 2 番目のビットストリームのコード化パラメータのうちの 1 つ以上が最初のビットストリームのコード化パラメータと違うの大多数を持っている圧縮されたイメージデータの 2 番目のビットストリームのための 2 番目の動作情報を補外するために、格納された最初の機能ポイント動作関連コード変換ヒントメタデータを利用するステップと、

前記 2 番目のビットストリームのアウトプットを提供するステップと

を含むことを特徴とするコード変換方法。

【請求項 5 2】 動作メタデータ、テクスチャー／エッジメタデータ、機能ポイント、および関連した動作メタデータ、特に、クレーム毎の新しい機能ポイントの数を使ってソースビデオのコンテンツを分類するステップ

を含むことを特徴とする分類方法。

【請求項 5 3】 転送コードのコード変換パラメータ設定を決定するために

使われる

ことを特徴とする請求項 5 2 記載の分類方法。

【請求項 5 4】 前記ビデオ分類に基づく視聴覚素材を組織するために使われる

ことを特徴とする請求項 5 2 記載の分類方法。

~~【請求項 5 5】 請求項 1 の方法を適用する装置からなる、入力されたビデオ／オーディオシグナルを処理する~~

ことを特徴とするビデオ／オーディオシグナル処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、ビデオ／オーディオシグナル処理方法およびビデオ／オーディオシグナル処理装置と関連し、マルチメディア記述とコード変換ヒントメタデータのコンパクトな表現を取り出す効率的な方法を提供する。

【0002】

コード変換ヒントメタデータは、種々の（例えばMPEG）圧縮されたコンテンツ表現とフレームレート、ビットレート、セッションのサイズ、量子化パラメータ、およびピクチャコードタイプ構造(GOP)のようなビットストリーム（例えばMPEG）パラメータのマニピュレーションの間でコード化している転送を容易にする。

【0003】

この発明のより一層のアプリケーションは、視聴覚分類とマルチメディア情報検索にある。

【0004】

【従来の技術】

視聴覚コンテンツの増大率は、現在、多くは光学式で、ワイアレスまたはワイヤードでつながっているネットワークの上に送られる。（図1）。

【0005】

これらのネットワークが種々のネットワークバンド幅強制するものにより特徴

付けられているので、要件は、変化主体ビジュアル品質を結果として生じる違うビットレートによって視聴覚コンテンツを表すように出現する。視聴覚コンテンツの圧縮された表現の付加的な要件は、視聴覚ターミナルのコンピュータ処理の機能だけでなくスクリーンサイズ、およびメモリー強制するものによって課される。

【0006】

従って圧縮されたフォーマットそれゆえの視聴覚コンテンツの貯蔵可能MPEG (MovingPictureExpertGroup) により定義されて種々のビットレート、フレームレート、スクリーンサイズに変換され、および様々な視聴覚ターミナルの解読複雑さとメモリー強制するものを変えるために。

【0007】

種々のネットワークバンド幅および種々の視聴覚ターミナルのための多い様々な圧縮された表現において視聴覚コンテンツの記憶装置を避けるために、蓄えられた圧縮されたMPEGフォーマットから、違うMPEGフォーマットにコード化している低い遅延およびリアルタイム転送が必要である。

【0008】

従って、抑制されるコンピュータ処理複雑でない低い遅延とリアルタイム実行は、コード変換方法の1番目の要件である。

【0009】

コード変換方法の2番目の要件は主体の視聴覚品質をなるべく保存する必要がある。

【0010】

コード変換方法を容易にし、これらを果たす様々な圧縮された目標フォーマットのための両方の要件ためには、コード化ヒントメタデータを事前に生成して転送し、それに圧縮された視聴覚コンテンツと一緒に、または別々に供給する。

【0011】

記憶装置サイズを減らし、マルチメディア記述とコード変換ヒントメタデータの配布 (例えば、ローカルな視聴覚コンテンツサーバーに放送してください) を容易にするように、非常にコンパクトな表現が必要である。

【 0 0 1 2 】

視聴覚コード変換ヒントメタデータのコンパクトな表現だけでなく自動的な抽出方法、および機器は、この発明において明らかにされる。

【 0 0 1 3 】

ビデオコード変換に関連した文献の存在：

~~In-W00-9-8-3-8-8-0-0A1: 0: H. Werner, N. D. Wells, M. J. Knee: Digital~~
Compression Encoding with improved quantization, 1998 (1つの順応的な量子化手法を持つデジタルの圧縮符号化が提案される。)

US5870146: Zhu Qin-Fan: Device and method for digital video transfer coding, 1999.

W009929113A1: Nilsson, Michael, Erling Ghanbari, Mohammed: Transfer coding, 1999.

US5805224: Keesman Gerrit J, Van Otterloo Petrus J.: Method and Device for Transfer coding Video Signals, 1998.

W009943162A1: Golin, Stuart, Jay: Motion vector extrapolation for transfer coding video sequences, 1999.

US5838664: Polomski Mark D.: Video teleconferencing system with digital transfer coding, 1998

W009957673A2: Bailleul, Nicolas: Transfer coding of a data stream, 1999.

US5808570: Bakhmutsky Michael: Device and Method for pair-matching Huffman-Transfer coding and high performance variable length deencoder with two-word bit stream segmentation which utilizes the same, 1998

W009905870A2: Lemaguet, Yann: Method of Switching between Video Sequences and corresponding Device, 1999.

W009923560A1: LUDWIG, Lester BROWN, William YUL, Inn, J. VUONG, Anh, T. , VANDERLIPPE, Richard BURNETT, Gerald LAUWERS, Chris LUI, Richard APPLEBAUM, Daniel: Scalable networked multimedia

system and application, 1999.

しかし、ビデオコード変換のための以前に言及された関連した背景技術は、視聴覚コード変換を容易にするように、コード変換ヒントメタデータ情報を使用しません。

【0014】

鈴木輝彦：「マルチメディア情報検索システム、検索方法、マルチメディアフォーマット変換機器、および方法」。

【0015】

本発明者は、鈴木輝彦氏との議論によっていくつかのコード変換ヒントメタデータを得た。また彼は彼のアイデアに関する特許を同時に提出している。ソースフォーマットの1つの大型のブロックごとに256ビットを使って再符号化メタデータを提供するデータの組を再コード化しているMPEG-2によるテレビのためのSMPTE（映画とテレビエンジニアの社会）提案された標準に、関連した背景技術がある。

【0016】

しかし、これは抽出において背景技術を関連させて、コード変換ヒントメタデータの表現は、考察された目標アプリケーションのためのいくつかの不利を持っている。

【0017】

はじめに、コード変換ヒントメタデータ（GOP構造、量子化設定、動きベクトルなどのような）は、すべての単一のフレームのためのこの前の背景技術および視聴覚ソースコンテンツの大型のブロック（MB）に取り出される。この方法は、提出詳細であること、および満足している順応的なコード変換主体の視聴覚品質を広く保存しているコード変換策略をほのめかし、容易にするの有利を提供する。

【0018】

しかし、この前の背景技術のコード変換ヒントメタデータのサイズは非常に大きい。前の背景技術の1つの具体的なインプリメンテーションにおいて、256ビットのコード変換ヒントメタデータはMPEGビデオの大型のブロック（MB）で格

納される。この莫大のコード変換ヒントメタデータはそれに適当ではない。例えば、ローカルな（例えばホーム）視聴覚への放送された配布は、サーバーを満足させます。従って、コード変換ヒントメタデータのこの前の背景技術は、放送されたスタジオアプリケーションに限定される。

【0019】

~~よりさらに、より多く、およびコード変換ヒントメタデータ抽出と表現の他の~~
前の背景技術は、具体的なビットレートによって、他の圧縮されたフォーマットとビットレートに、圧縮された視聴覚ソースコンテンツのコード変換のための一般コード変換ヒントメタデータを収集する。

【0020】

しかし、この前の背景技術は、コード変換コンテンツの独特な特性を考慮しないことの不利を示す。

【0021】

例えば、ソースコンテンツにおいて、視聴覚特徴が、動作の制限された量および小さい詳細（例えばニュースアナウンサー場面）から、速く描いている他の視聴覚セグメント（例えばスポーツの場面）に視聴覚セグメントが変更する動作。

【0022】

この他の前の背景技術は、欺きコード変換ヒントメタデータを選び、両方のビデオセグメントの種々の特徴を適切に表さず、従って、不利なビットレート割り当てによって不利な視聴覚品質に結果として生じています。

【0023】

【発明が解決しようとする課題】

前述の最先端を考慮して、それは、抽出方法と装置をコンパクトに提供する現在の発明のオブジェクトであるけれども、視聴覚は、順応的なマルチメディア記述とコード変換ヒントメタデータ表現を満足させます。

【0024】

この発明は、サポートコード変換メタデータを使って、1つの圧縮されたフォーマットからの圧縮された視聴覚コンテンツから他のフォーマットの視聴覚コンテンツにコード化している転送のフィールドにある。

【 0 0 2 5 】

用語コード変換は含むけれども、圧縮されたフォーマット（例えば、MPEG-2 フォーマットからMPEG-4 フォーマットへの変換）、フレームレート変換、ビットレート変換、セッションのサイズ変換、スクリーンサイズ変換、ピクチャコード化タイプ変換などを変更することに制限されませんコード変換システムの主要な強制するものは下記である。

【 0 0 2 6 】

- 1.) コード変換プロセスを通して視聴覚品質をなるべく保存する、
- 2.) およびリアルタイムアプリケーションとを低く可能にするためにコンピュータ処理の複雑さを最小遅延となるように制限する。

【 0 0 2 7 】

これらの強制するものを達成するために、コード変換ヒントをカバーしている付加的なデータ（メタデータ）が、圧縮された視聴覚コンテンツに関連できます。

【 0 0 2 8 】

この発明の他のアプリケーションは、ビデオの違う場面活動のクラスとして、言及されたコード変換ヒントステートを使って自動的なビデオ分類にある。

【 0 0 2 9 】

【課題を解決するための手段】

本発明のビデオ／オーディオシグナル処理方法は、ビデオ／オーディオシグナルを供給し、処理するビデオ／オーディオシグナル処理方法であって、コード変換目標ビットストリームパラメータを記述するステップと、コード変換ヒントメタデータを抽出するステップと、コード変換ヒントメタデータを記憶するステップと、セグメントへの視聴覚素材を分離するステップと、視聴覚セグメントに対してコード変換ヒントメタデータを関連させるステップと、コード変換するステップとを含むことを特徴とする。

【 0 0 3 0 】

【発明の実施の形態】

図 1 は、家庭のネットワーク環境においてコード化している転送のための典型

的なシナリオにおいて一般概要を示す。

【0031】

視聴覚コンテンツサーバーはユニット102において描かれて、視聴覚コンテンツ記憶装置103、視聴覚コード変換ユニット106、コード変換ヒントメタデータ抽出ユニット104、および視聴覚コード変換ヒントメタデータ記憶装置105を含んでいます。

【0032】

視聴覚コンテンツ記憶装置103は、変化ビットレートおよび主体の品質を変えることによって、様々なソースから、圧縮された視聴覚素材を含んでいます。例えば、視聴覚コンテンツ記憶装置は、非常に低いビットレートによって、ポータブルなDVビデオカメラ111、および約MPEG-4圧縮ビデオから、ホーム・ビデオを、MPEG-4インターネットカメラ112、すでにコード変換ヒントメタデータと場合によっては関連している101からの約5 Mbit/sのMPEG-2 MP@ML圧縮放送されたビデオからの約10 kbit/sに含んでいるかもしれません。

【0033】

視聴覚コンテンツサーバー102は、高い定義圧縮MPEGビデオをかなりより高いビットレートに含んでいたほうがよいでしょう。

【0034】

視聴覚コンテンツサーバーはネットワークに接続し、典型的である。ワイヤに基づくまたはワイアレスな家庭ネットワーク113。

【0035】

このネットワーク113に、例えば種々の特徴を持ついくつかの視聴覚ターミナルが取り付けられる。ワイアレスなMPEG-4視聴覚PDA107、高精細テレビエンターテインメント108、テレビゲームコンソール109、およびITU-T (国際的なテレコミュニケーション連合技術標準グループ) H.263のための高解像度視聴覚ターミナルは、テレビ電話110の基礎に相当していました。

視聴覚ターミナルは、種々のビットレートトランスミッション機能(ケーブルまたはラジオリンクによる)によって家庭のネットワークに取り付けられる。

例えば、ワイアレスなビデオPDA (パーソナル・デジタル・アシスタント) 10

7はこのネットワークに接続し、それは、コンピュータ処理のパワー、記憶装置、スクリーンサイズ、ビデオフレームレート、およびネットワークビットレートについて厳しく制限できます。従って、視聴覚コード変換ユニットは例えば、コードを転送する必要があり、視聴覚に含まれている720x480ピクセルは、ヨーロッパ25fps（フレーム/秒）の5Mbit/sMPEG-2放送されたビデオワイアレスなトランスミッションのためのMPEG-4-5.00kbit/s-15fpsビデオ、およびワイアレスなMPEG-4ビデオPDA107の近くの352x240ピクセルディスプレイにおけるディスプレイにサーバー102を満足させます。

【0036】

視聴覚コード変換ユニット106はコード変換ヒントメタデータ105を使い107、108、および109、または110、リアルタイムのコードを転送するために、視聴覚の圧縮されたソースビットレートは、具体的な目標視聴覚ターミナルの機能に満足させます。コード変換ヒントメタデータは、コード変換ヒントメタデータ抽出ユニット104において生成されるか、またはそれらが、放送されたサービス101により配布される。

【0037】

視聴覚コンテンツバッファ103から、視聴覚コード変換ユニット106、ソースフォーマットの圧縮されたビットストリーム116が与えられる。このビットストリームはまた、以下において最初のビットストリーム116として示される。目標フォーマット115（またこの2番目のビットストリームとして示された）のビットストリームは、106において家庭のネットワーク113にコード化している転送の後に与えられる。家庭のネットワークから、視聴覚コンテンツ記憶装置103において、114により、中のコンテンツ、例えば圧縮されたDVフォーマットが蓄えられる。

【0038】

コード化ヒントメタデータ抽出を転送

図2は、コード変換ヒント抽出、コード変換ヒント記憶装置、およびコード変換プロセスにおいて、一般化した概要を示す。

【0039】

バッファ201は視聴覚コンテンツをソースフォーマットで含んでいます。

【0040】

バッファ202は、ソースフォーマット、ように織り混ざったなどの記述を含んでいますまたは進歩的なフォーマットなどバッファ203は、目標フォーマット、ように織り混ざったなどの記述を含んでいますまたは進歩的なフォーマットなどコード変換ヒント抽出ユニット207は、視聴覚バッファから、コンテンツを、圧縮されたソースフォーマットで読み、バッファ202からのソースフォーマット記述だけでなく201、およびコード変換目標は、203から記述をフォーマットする。

【0041】

207のコード変換ヒントを計算した後に、コード変換ヒントはコード変換ヒントメタデータバッファ206に格納される。

【0042】

視聴覚コード変換ユニット205は、206において格納されたコード変換ヒントメタデータによってソースフォーマット視聴覚コンテンツバッファ201、および変化から、目標フォーマットにソースフォーマットを読みます。

【0043】

視聴覚コード変換ユニットのアウトプット205は、視聴覚目標フォーマットバッファ209の新しい圧縮されたフォーマットで蓄えられる。

【0044】

204は、コンテンツ記憶装置からソースフォーマット201に視聴覚コード変換ユニット205に最初のビットストリームを説明する。

【0045】

208は、視聴覚コード変換ユニット205から視聴覚目標フォーマットバッファ209に2番目のビットストリームを説明する。

【0046】

【表 1】

State	M	bitrate_fraction_for_I	bitrate_fraction_for_P
1	5	0.8	0.15
2	4	0.85	0.1
3	3	0.9	0.05

図3と図4は、実世界の例を使ってコード変換ヒントメタデータ組織という基本的な原則を示す。一般的に言って、MPEGに基づくビデオ圧縮により、連続したフレームの間の変化だけが符号化される予測方法が使われる。1つのフレームから次のフレームに高い量の変化に満足しているビデオが種々の再符号化パラメータ設定を必要とすること（ビットレートを制限する間主体の品質を運営するために）、ビデオよりフレームの間の小さな変更~~に満足している。~~従って、再符号化パラメータにおいて事前に決めることが重要である。

【0048】

従って、コード変換ヒントメタデータ選択は主に、予測不可能なビジュアルなコンテンツの量と特徴に依存する。この新しいビジュアルなコンテンツは、前のフレームから予測できず、ビットレート集中的符号化された使用DCT-係数である必要がある。従って、明らかにされた方法は、新しい機能ポイントの数を使い、それは、フレームあたり新しいコンテンツの量を決定するために、前のフレームから現在のフレームに追跡されない。

【0049】

図3は、ビデオのフレーム番号に依存しているフレームあたり新しい機能ポイントの数のグラフを示す（水平の軸、時間軸）。

【0050】

セクション301はビデオセグメントの一部を説明し、次のフレームの間で非常に少ない多くの新しいコンテンツだけが出現し、従ってそれぞれのコード変換ヒントメタデータ（例えば、大きいGOPサイズ低いフレームレート低ビットレート、…）は選ばれます。

【0051】

セクション302は、意味しているフレームあたり新しい機能ポイントの少し高い数~~を示し、~~国の説明コード変換ヒントメタデータは選ばれて、それは最適なコード変換パラメータをこの状況に提供する（例えば、少し小さいGOPサイズより高いビットレート）。

【0052】

セクション303はコード変換メタデータを示す場面あたりフレームあたり新

しい機能ポイントの高い数と従って高い量の新しいコンテンツというステートをほのめかす。従って、より小さいM値 (I/P-フレーム距離) およびより高いビットレートは選ばれる必要がある。

【 0 0 5 3 】

図4は、3つの離散的なコード変換ヒントメタデータステートから成っているコード変換ヒントメタデータステート図の基本的な組織の例を示す。

【 0 0 5 4 】

すべての離散的なコード変換ステートは、GOP構造、量子化パラメータ、ビットレート、スクリーンサイズなどのためのメタデータを含んでいますこれらのコード変換ヒントパラメータは、固定されている値を持っているか、または他のパラメータの機能であるかもしれない。

【 0 0 5 5 】

例えば、GOP長さは、フレームあたり新しい機能ポイントの数の離散的な機能であるかもしれず、量子化パラメータは、DCT係数から引き出されたエッジとテクスチャー活動の機能であるかもしれない。

【 0 0 5 6 】

3つの種々の符号化状況を適応させるために、この例の3つのコード変換ヒントメタデータステートのうちのそれぞれが選ばれる。この図において、ステート3と403は、高い量の動作、およびフレームあたり低い量の新しいコンテンツのために選ばれて、このステートのための最適なコード変換ヒントメタデータを表す。

【 0 0 5 7 】

ステート2と402が、選ばれた低い量の動作であること、および高い量のコンテンツと費やされるビットの高い数を必要とする活動を高く研ぐ。例えば、ステート1と401は、低い場面活動のためのコード変換プロセスを適応させるために選ばれる。ビデオ編集効果急場面変化 (違う相互のフェード効果のような) または黒に提供された他の特別なコード変換ヒントメタデータステートもある。2つの場面の間で描く。ビデオ編集効果の位置は半自動に、または十分に、自動的に手動で検出できる。

【 0 0 5 8 】

図 5 は、コード変換ヒントメタデータ抽出において、圧縮されたおよび未圧縮であったソースコンテンツから概要を示す。

【 0 0 5 9 】

5 0 1 が視聴覚ソースコンテンツバッファであること、および 5 0 2 ソースフォーマット記述バッファおよび 5-0-3 目標フォーマット記述バッファ。

【 0 0 6 0 】

5 0 4 は、圧縮されたまたは未圧縮であった領域から、動きベクトル、DCT-係数、および機能ポイント抽出を含んでいます。これテクニックは背景技術での熟練のために知られて、従って、簡単に要約されるだけである。

【 0 0 6 1 】

圧縮された領域では、P および B マクロブロックからの動きベクトルはビットストリームから直接取り出される。しかし、マクロブロックに動きベクトルが全然ありません。

【 0 0 6 2 】

従って、何人かの著者 (Roy Wang, Thomas Huang: "Fast Camera motion Analysis in MPEG domain", IEEE International Conference on Image Processing, ICIP 99, Kobe, Japan, Oct 1999) は、I マクロブロックのための B および P マクロブロックのために得られた動きベクトルを補間する。

【 0 0 6 3 】

マクロブロックのブロックのための DCT 係数はビットストリームから直接取り出される。P および B マクロブロック限られた数 DCT-係数 (DC および 2 つの AC 係数) のために、Shih-Fu チャン、デイビッド G により説明された方法により得られえます (Shih-Fu Chang, David G. Messerschmid: "Manipulation and Composition of MC-DCT compressed video", IEEE Journal on Selected Areas in Communications, vol. 8, 1996.)。

【 0 0 6 4 】

圧縮された領域機能ポイント抽出と動作評価は、本発明者の特許出願により明らかにされる。(「圧縮された領域機能のための方法と装置」、PCT 特許出願、

1 9 9 9 年 1 2 月 出 願) 。

【 0 0 6 5 】

場合によっては、視聴覚ソースコンテンツは、未圧縮であったフォーマット、またはDCTおよびMPEG-1、MPEG-2、MPEG-4、ITU-TH.261、およびITU-TH.263で使用される動作補償原則に基づかない圧縮フォーマットで入手可能であるだけである。

【 0 0 6 6 】

DVフォーマットのために、DCT-係数だけが入手可能である。これらのケース動きベクトルにおいて、背景技術、例えば本発明者の熟練に既知の動作評価方法により得られる。: "Algorithms, Complexity Analysis and VLSI Architectures for MPEG-4 Motion Estimation", Kluwer Academic Publishers, 1999。

【 0 0 6 7 】

DCT-係数は、ブロックに基づくDCT-変化を実行することによって得られる。(K.R. Rao, P. Yip: "Discrete Cosine Transform - Algorithms, Advantages, Applications", Academic Press 1990)。

【 0 0 6 8 】

例えば、ピクセル領域(未圧縮であった領域)についての機能目的は、Bruce D. Lucas, Takeo Kanade: "An iterative registration technique with an application to stereo vision", International Joint Conference on Artificial Intelligence, pp674-679, 1981により説明された方法により得られる。

【 0 0 6 9 】

505は、パラメータの動作モデルのパラメータが504の動きベクトル表現から取り出される動作分析部分を説明する。6と8つのパラメータを持つパラメータの動作モデルは共通であり、パラメータの動作評価は背景技術で既知である。(M. Tekalp: "Digital Video Processing", Prentice Hall, 1995.)

ここの動作表現を使うゴールは、遅延とスピード理由のため、変化エンコーダでの動作評価を取り除くことである。従って、ソースビットストリームからの動作のインプット表現は、アウトプット表現を引き出すために使用される(目標ビッ

トストリーム)。例えば、動作表現においてリサイズ、組み合わせ進歩的な変換信頼を重くスクリーン大きさで分ける。動作表現のパラメータは、また、GOP構造についての決定をコード化するために使われます。

【 0 0 7 0 】

5 0 6 は、ビットストリームから取り出された DCT-係数に基づくテクスチャー／エッジ分析部分を示す。~~(K.R. Rao, P. Yip: "Discrete Cosine Transform - Algorithms, Advantages, Applications", Academic Press 1 9 9 0, or K.W. Chun, K.W. Lim, H. D. Cho, J.B. Ra: "An adaptive perceptual quantization algorithm for video encoding, IEEE Transactions on Consumer Electronics, Vol. 3 9, No. 3, August 1 9 9 3)~~

5 0 7 は、圧縮された領域のための機能ポイント追跡部分を参照する。使用されたテクニックは、先願において説明されている：Peter Kuhn: "Method and Apparatus for compressed domain feature point registration and motion estimation", PCT patent, December 1 9 9 9.

5 0 8 は新しい機能の番号を計算する。フレーム、言及された特許において明らかにされるテクニックで指し示す。

【 0 0 7 1 】

5 0 9 は時間的なビデオセグメンテーションを計算し、5 1 0 はコード変換ヒントを計算する。すべてのセグメントのために述べる。これらのテクニックの適切な方法は、図 6、図 7、および図 8 に説明される。

【 0 0 7 2 】

5 1 1 は、動作関連コード変換ヒントメタデータを含み、それは図 1 5 において説明される。5 1 2 は、テクスチャー／エッジ関連コード変換ヒントメタデータを含み、それは図 1 5 において説明される。

【 0 0 7 3 】

5 1 3 は、機能ポイントコード変換ヒントメタデータを含み、それは図 1 5 において説明される。5 1 4 は、ビデオセグメントコード変換ヒント選択メタデータを含み、それは図 1 6 において説明される。次のセクションは、コード変換ヒ

ントメタデータの自動的な抽出、コンパクトな表現、および用法を説明する。

【0074】

図6はビデオセグメンテーションを明らかにし、コード変換ヒントは選択方法を述べます。601において、いくつかの変数が初期設定される。可変の「フレーム」はソースビットストリームの現在のフレーム番号であり、「nフレーム」は新しいビデオセグメント内のフレームの数である。(または、GOPグループのピクチャ数)。

【0075】

他の変数はこのルーチン内の使用をもっているだけである。602において、このGOP内のフレームの数がインクリメントされる。603 (図7) において、このフレーム内で新しいセグメント/GOPが始まるかどうか決定される。

【0076】

この時には、ケースステップ604には入り、さもないければ、ステップ615には入ります。604において、可変のlast_gop_startは「new_gop_start」値によって初期設定される。608と609において、1より「フレーム」変数が大きいといけなから、可変の「last_gop_stop」は、「frame-1」に設定される。さもないければ、1まで610の「last_gop_stop」に設定される。

【0077】

次のステップ、図8において詳細に描かれる611は、動作パラメータ605、テクスチャー/エッジパラメータ606、および機能ポイントデータ607に依存するコード変換ヒントステートを決定する。

【0078】

ステップ612において、コード変換ヒントメタデータはコード変換ヒントメタデータバッファに出力される。この例に、コード変換ヒントメタデータが「nフレーム」(このGOP内のフレームの数)から成っているコード変換ヒントは、新しいGOPのすべてのパラメータと開始フレーム数によって述べます(new_gop_start)。

【0079】

その後に、可変の「nフレーム」は、0および「機構」が可変の「new_gop_s

「tart」に与えられる現在のフレーム数にセットされる。その時、ステップ615において、もしソースビットストリームのすべてのフレームが処理されたならば、それはテストされる。

【0080】

もしこれがケースではないならば、ステップ614は入力されてどの隣のフレーム数へのインクリメント、ステップ602によって再スタートする。アルゴリズム。さもなければ、アルゴリズムは停止する。

【0081】

図7は、もし可変の「nフレーム」がMの整数倍数であるならば（I/Pフレーム距離である）図6から、開始フレームおよび新しいビデオセグメントまたはGOP701の終わりフレームを決定するいくつかの可能な方法のうちの1つを明らかにする。

【0082】

この場合は真であり、その時、「いいえ」が選ばれて、702において、もし現在のフレーム数が最初のフレームであるならば、それはテストされる。

【0083】

これが真実であるならば、ステップ703は選ばれて、もしGOP内の「gop_min」というフレームの最小の数より「nフレーム」が大きいならば、それはテストされる。

【0084】

702の結果が「はい」であるならば、新しいGOPは705内で始められる。703の結果がはいであるならば、新しいGOPは705に取りかかります。

【0085】

703の結果が「いいえ」であるならば、テスト704には入ります。704の結果が「はい」であるならば、GOPはステップ706内で閉じられて、さもなければ、このサブルーチンは終了する。

【0086】

図8は、フレームあたり新しい機能ポイントの数だけを考慮している具体的なGOPまたは視聴覚セグメントのためのコード変換ヒントステートの選択のための

例を明らかにする。

【0087】

この基本的なアイデア類似決定に基づき、構造化し、DCT-係数から得られたテクスチャー／エッジパラメータだけでなくパラメータの動作評価から、前記する動作パラメータを使うので背景技術で熟練したもののため実装できることに注意せよ。

【0088】

ここでアルゴリズムのこのクラスが明らかにしたことは、また、動作、エッジ活動、視聴覚分類のより高いレベルをもたらしているフレームあたり新しいコンテンツなどについて視聴覚素材を分類するために用いられる。

【0089】

このケースにおいて、コード変換ヒントステートはその時種々のコンテンツ素材の具体的なクラスを表す。

【0090】

801において、このサブルーチンの変数が初期設定される。可変の「frame_no」は、「last_gop_start」パラメータのコンテンツを与えられて、可変の「合計」は0によって初期設定される。その時は、ステップ802、「合計」変数のコンテンツにおいて現在のフレーム、「frame_no」の新しい機能ポイントの数によりインクリメントされる。

【0091】

ステップ803において、テストされ、もし可変の「last_gop_stop」のコンテンツより可変の「frame_no」のコンテンツがより小さいならば、ステップ804へ進入し、さもなければ、再びステップ802には入る。

【0092】

ステップ804において、もし前もって決定されたパラメータ「summax」のうちの1つの8つより可変の「合計」の値が小さいならば、それは、この例の値200を持っていてテストされる。

【0093】

もしこれが真実ならば、コード変換ヒントステート1（ステップ806）が選

ばれて、そのためにパラメータは図8のテーブル内で描かれます。

【0094】

ステップ805において、もし前もって決定されたパラメータ「summax」の1つの4分の1より可変の「合計」の値が小さいならば、それはテストされる。

【0095】

もしこれが真実ならば、コード変換ヒントステート2(ステップ807)が選ばれます。もしこれが条件であるならば、間違いで、コード変換ヒントステート3(ステップ808)に入り、サブルーチンは終了する。

【0096】

804と805の決定入り口はコード変換ヒントステートの定義と数に依存することに注意してください。

【0097】

3. 転送コード化ヒントメタデータに関する記述

メタデータ説明のために、仮のC-コードスタイルが使われます。記述のための短縮Dおよび記述体系のためのDSは出現MPEG-7メタデータ標準において定義されるように使われます。

【0098】

コード変換ヒントメタデータの構造の組織

図9は、一般的な視聴覚記述体系内のコード変換ヒントメタデータの構造の組織を示す。901が一般的な視聴覚DSすなわちセグメントDS904、およびメディア情報DS902を説明する。

【0099】

セグメントDS904から、セグメント分解906が引き出される。

【0100】

セグメント分解906から、ビデオセグメントDS907、および動き領域DS907が引き出される。

【0101】

ビデオセグメントDS907から、セグメント基づいたコード変換ヒントDS909(図14)が引き出される。

【0102】

ビデオセグメントDS907の通路それまたはいくつかのコード変換ヒントステートDS911（図16）。

【0103】

動き領域DS908から、領域を動かすためのセグメント基づいたコード変換ヒントDS910（図14）が引き出されて、それは、それまたはいくつかのコード変換ヒントステートDS912、および図16にアクセスする。

【0104】

メディア情報DS902から、メディアプロフィールDS903が引き出される。メディアプロフィールDS903から、一般コード変換ヒントDS905、（図10）が引き出される。

【0105】

図10は、コード変換ヒントDS1001、ソースフォーマット定義DS1002（図11）、およびまたは目標フォーマット定義DS1003、および図12のいくつかの例の1つの例から成っている構造の組織を示す。

【0106】

付加的に、コード変換ヒントDS1001は、一般コード変換ヒントDS1004（図13）、および1つのオプションのコード変換符号化複雑度DS1005（図15）の1つのオプションの例から成っています。

【0107】

転送コード化ソースフォーマットの定義

図11は、ソースフォーマット定義コード変換ヒントメタデータを示し、それは、全体の視聴覚コンテンツまたは具体的な視聴覚セグメント、および図9に関連する。以下において、適切な記述子と記述体系が明らかにされる。

ビットレートはタイプをもち(整数)、ソース視聴覚データストリームの2番目あたりビットレートを説明する。

【0108】

size_of_picturesはタイプをもち<2*int>、xとy方向のソース視聴覚フォーマットのピクチャのサイズを説明する

number_of_frames_per_secondはタイプをもち(整数)、ソースコンテンツの2番目あたりフレームの数を説明する。

【0109】

pixel_aspect_ratioはタイプをもち<float>、ピクセル縦横比を説明する。

【0110】

pixel_colour_depthはタイプをもち(整数)、色深さを説明する。

【0111】

usage_of_progressive_interlaced_formatはサイズをもち<1 bit>、進歩的であることによるソースフォーマットの申し込みまたは織り混ぜられたフォーマットで説明する。

【0112】

usage_of_frame_field_picturesはサイズをもち<1 bit>、もしフレームまたはフィールドピクチャが使われるならば説明する。

【0113】

compression_methodはタイプをもち(整数)、ソースフォーマットのために使われた圧縮方法を定義し、リストのうちのものである。: MPEG-1、MPEG-2、MPEG-4、DV、H.263、H.261などすべての圧縮方法より一層のパラメータのために、ここで定義される。

【0114】

GOP_structureは、I/P/B-ステートのランレングス符号化データフィールドである。例えば、MPEG-2ビデオにだけI-フレームがあるならば、圧縮された領域のDVフォーマットへの直接的な変換が可能である。

【0115】

転送コード化目標フォーマットの定義

図12は、目標フォーマット定義コード変換ヒントメタデータを示し、それは全体の視聴覚コンテンツまたは具体的な視聴覚セグメントに関連できる。以下において、適切な記述子と記述体系が明らかにされる。

【0116】

ビットレートはタイプをもち(整数)、目標視聴覚データストリームの2番目あ

たりビットレートを説明する。

【0117】

size_of_picturesはタイプをもち<2*int>、xとy方向の目標視聴覚フォーマットのピクチャのサイズを説明する。

【0118】

number_of_frames_per_secondはタイプをもち(整数)、目標コンテンツの2番目あたりフレームの数を説明する。

【0119】

pixel_aspect_ratioはタイプをもち<float>、ピクセル縦横比を説明する。

【0120】

pixel_colour_depthはタイプをもち(整数)、色深さを説明する。

【0121】

usage_of_progressive_interlaced_formatはサイズをもち<1 bit>、もし目標フォーマットが進歩的か、または織り混ざる必要があるならば説明する。

【0122】

usage_of_frame_field_picturesはサイズをもち<1 bit>、もしフレームまたはフィールドピクチャが使われるならば説明する。

compression_methodはタイプをもち(整数)、ソースフォーマットのために使われた圧縮方法を定義し、リストのうちのものである。: MPEG-1、MPEG-2、MPEG-4、DV、H.263、H.261などすべての圧縮方法より一層のパラメータのために、ここで定義される。

【0123】

GOP_structureは、I、P、B-ステートのオプションのランレングス符号化データフィールドである。このオプションのパラメータによって、固定されているGOP構造は強制できます。

【0124】

例えば、固定されているGOP構造は、一定の位置で、ビデオ編集を容易にすることをI-フレームに強制するために有益であるかもしれません。

【0125】

一般コード変換ヒントメタデータ

図13は一般コード変換ヒントメタデータを示し、それは全体の視聴覚コンテンツまたは具体的な視聴覚セグメントに関連できます。以下において、適切な記述子と記述体系が明らかにされる。

【0126】

`use_region_of_interest_DS`はその長さを持ち、 $\langle 1\text{-bit} \rangle$ 興味記述体系の領域がコード変換ヒントとして入手可能であるかどうかを示す。

【0127】

`region_of_interest_DS`が使われるといけないからその時例えば以下のうちの1つであるかもしれない`shape_D`(: 興味の領域を空間、時間的に説明するために、`boundary_box_D`、`MB_shape_D`、または`motion_trajectory_D`とともにどのような他の`shape_D`)でも使用できます。MB形記述子は、オブジェクト形記述のための大型のブロック(16x16)寸法決めブロックを使います。

【0128】

`motion_trajectory_D`は、`start_frame`およびこの`region_of_interest_DS`の`end_frame`が定義されるように、すでに時間の観念を含みます。`region_of_interest_DS`は個々の`shape_D`および個々の`motion_trajectory_D`のサイズを持っています。コード変換アプリケーションのために、`region_of_interest_DS`が使われえて、例えば、バックグラウンドにわたってより多いビット(または、量子化をそれぞれ修正してください)を興味の領域内のブロックに使うために。

【0129】

MPEG-4への他の転送コード化アプリケーションが、seperate MPEG-4 オブジェクトによって興味の領域を説明するし、バックグラウンドのような他のMPEG-4オブジェクトより興味の領域のためのより高いビット速度およびより高いフレームレートに費やされるために。

【0130】

この`region_of_interest_DS`の抽出は自動的に、または手動で実行できます。

【0131】

`use_editing_effects_transfercoding_hints`はその長さを持ち $\langle 1\text{ bit} \rangle$ 、編集

効果に基づくコード変換ヒントで情報が入手可能であるかどうかを示す。

【0 1 3 2】

camera_flashは、すべての入力、カメラフラッシュが起こるフレーム数を説明するいくつかの入力のリストである。従って、記述子の長さは、整数(整数)で増加したカメラフラッシュイベントの数である。コード変換アプリケーションのために、カメラフラッシュ記述子は、/ビデオ再エンコーダのほとんどとしてエンコーダ使用を転送し輝度違い、非常に有益である。Peter Kuhn: "Algorithms, Complexity Analysis and VLSI Architectures for MPEG-4 motion estimation", Kluwer Academic Publishers, 1 9 9 9 .

輝度に基づく動作評価の場合には、予測のために、2つの次のフレーム（フラッシュ、フラッシュなしのものを持つもの）の2つの大型のブロックの間の平均絶対物エラーが高すぎて、カメラフラッシュを持つフレームは、高いビットレートコストによって形成内として符号化される必要がある。

【0 1 3 3】

従って、コード変換内のカメラフラッシュを示すことにより、例えば記述体系が、適量なビットコストによって、使用に、輝度訂正動作評価方法または他の方法がカメラフラッシュによって錨フレーム (s) からフレームを予測することを可能にするのがほのめかされる。

【0 1 3 4】

cross_fadingは、すべての入力、相互のフェードの開始フレームと終わりフレームを説明するいくつかの入力のリストである。従って、この記述子の長さは相互のフェードイベントの数で2回（整数）である。

【0 1 3 5】

コード変換ヒントメタデータの相互のフェードイベントを示すことは、相互のフェードの間にビットレート/量子化をコントロールするために非常に有益である。相互のフェード予測の間に、ふつう、予測エラーコード化のためのビットレート増加を起こす制限された使用である。場面をフェードさせている十字の間に通常ぼやけると、ビットレート増加は、量子化スケール、ビットレート、またはレートコントロールパラメータをそれぞれ調整することによって制限できます。

【 0 1 3 6 】

`black_pictures`は、すべての入力が始動フレームおよび一連の黒いピクチャの終わりフレームを説明するいくつかの入力のリストである。場面の間では、特にホーム・ビデオにおいて、黒ピクチャは起こるかもしれません。実験的に、予測が、制限された使用をもっているだけであるので、示して`that`動作補いDCTエンコーダの一連の黒いピクチャ増加ビットレート結果として生じています。従って、このコード変換ヒント記述子は、黒いピクチャの間に、量子化`_scale`、ビットレート、またはレートコントロールパラメータをそれぞれ調整してビットレートを制限するために用いられます。

【 0 1 3 7 】

`fade_in`は相互のフェードと同様で、中でのフェードの開始フレームと終わりフレームを決定する入力の多くと評される。相互のフェードに比べて、黒いピクチャからの開始でのフェード、および従って目の一種のマスキング効果は、中でのフェードの間に、量子化`_scale`を調整してビットレートを制限するために使われてビットレートか、またはそれぞれ評価できますコントロールパラメータ。

【 0 1 3 8 】

場面の後に、一連の黒いピクチャが説明されるので、`fade_out`は、どんなに違いによって次第にはっきりするように同様である。

【 0 1 3 9 】

`abrupt_change`は(整数)、どこで薄れなしの突然の場面またはショット変化が出現するかを示して単一のフレーム多くのタイプのリストによって説明される。

【 0 1 4 0 】

例えば、これらのイベントは図 3 の非常に高いおよび鋭いピークにより示される。これらのピークは新しいカメラショットまたは場面の最初を示す。`abrupt_change`編集効果はフェード効果と対比するとである。

【 0 1 4 1 】

その時の2つのビデオセグメントの間の突然の変化が出現する時には、新しい視聴覚セグメントの詳細を適応させて、認めるために、人ビジュアル知覚には数ミリ秒が必要である。人の目のこの遅いこと効果は、例えば、ビットレートを減

らすか、または場面またはショットの突然の変化の後にビデオセグメントの最初のフレームのための量子化_scaleパラメータを修正するために、ビデオコード変換のために有益に使われえます。

【 0 1 4 2 】

use_motion_transfer<1 bit>coding_hints_DS: し、その長さで動作関連コード変換ヒントメタデータの使用を示す。

【 0 1 4 3 】

number_of_regionsは、以下の動作関連コード変換ヒントメタデータが有効な領域の数を示す。

【 0 1 4 4 】

<1 bit>もし領域が矩形か、または恣意的に形づくられるならば現在それが長さのフィールドにより示されるfor_every_region。領域が恣意的に具体化されるといけなから、領域記述子（成っていること例えば形記述子および動作弾道記述子の間の）が使われます。矩形の領域の場合には、矩形の領域のサイズが使われます。

【 0 1 4 5 】

この領域内の動作フィールドはパラメータの動作モデルにより説明されて、それはフレームのすべてのフレームまたは連続のためのいくつかのパラメータのため決定される。

【 0 1 4 6 】

ソースビデオの現実の動作のこの動作表現をコード化している転送のために、(re-) 符号化部分のコンピュータ処理複雑動作評価の検索エリアを制限するために使われうるけれども、またのために速さおよび効率的な不、織り混ぜた／織り混ぜられた（フレーム／フィールド）変換およびビデオ内の動作の量に依存しているGOP（グループのピクチャ）構造を決定する。動作表現はまた、ビデオのサイズ変換のために有益に使われえます。

【 0 1 4 7 】

セグメント基づいたコード変換ヒントメタデータ

図 1 4 は、セグメントに基づくコード変換ヒントメタデータを示す。これは、一定

の特徴を描く視聴覚セグメントのための (re-) enencoder/変化エンコーダ設定を決定するために使用できます。以下において、適切な記述子と記述体系が明らかにされる。

【 0 1 4 8 】

start_frameはサイズをもち(整数)、視聴覚セグメントのコード変換ヒントメタデータの最初のフレーム番号を説明する。

【 0 1 4 9 】

nフレームはサイズをもち(整数)、視聴覚セグメントの長さを説明する。

【 0 1 5 0 】

I_frame_locationは、この視聴覚内のI-フレームの位置を説明するいくつかの将来性が分割する以下を提出する。

【 0 1 5 1 】

select_one_out_of_the_followingはサイズをもち<2 bit>、以下のI-フレーム位置記述方法のうちの1つを選びます。

【 0 1 5 2 】

最もfirst_frameはサイズをもち<1 bit>、デフォルトI-フレーム位置である。これは意味し、最初のフレームだけがこの視聴覚セグメントのIntraフレームであり、使われますすっかり、より一層の予測のための錨この視聴覚セグメント内の他のフレームが、PまたはBフレームである。

【 0 1 5 3 】

list_of_framesは、この視聴覚セグメント内の形成内であるフレーム数のリストを与えます。これは、視聴覚セグメント内の形成内の位置を恣意的に説明するために可能にする。このリスト内のkフレームのために、この記述子のサイズが<k*int>である。

【 0 1 5 4 】

最もfirst_frame_and_every_k_framesはサイズをもち(整数)、このセグメント内の最初のフレームはIntraであり、kは視聴覚セグメント内のI-フレームの間を説明する。

【 0 1 5 5 】

no_I_frameがサイズをもち<1 bit>、この視聴覚セグメントの符号化が前のセグメントの錨（形成内）に基づく時には、有益でケースを説明し、I-フレームはこの視聴覚セグメント内で全然使われません。

【 0 1 5 6 】

量子化_scaleはサイズをもち(整数)、初期の量子化_scale値をこの視聴覚セグメントに記述する。

【 0 1 5 7 】

target_bitrateはサイズをもち(整数)、2番目あたり目標ビットレートをこの視聴覚セグメントに記述する。

【 0 1 5 8 】

target_max_bitrateはサイズをもち(整数)、2番目あたり最大の目標ビットレートをこの視聴覚セグメント（オプションの）に記述する。

【 0 1 5 9 】

target_min_bitrateはサイズをもち(整数)、2番目あたり最小の目標ビットレートをこの視聴覚セグメント（オプションの）に記述する。

【 0 1 6 0 】

use_transfer<1 bit>coding_statesはサイズをもち、この視聴覚セグメントのためにコード変換ヒントステートが使われるかどうかを記述する。

【 0 1 6 1 】

転送coding_state_nrはサイズをもち(整数)、このセグメントのためのコード変換ヒントメタデータステートを与えます。コード変換ヒントメタデータステートは、コード変換ヒントステートのテーブルの入力にポインタである。コード変換ヒントステートのテーブルはいくつかの入力を持ち、新規参入はコード変換ヒントパラメータにより追加されるか、または削除できます。単一のコード変換ヒントステートのコード変換ヒントメタデータは、図 1 6 において説明される。

【 0 1 6 2 】

add_new_transfer<1 bit>coding_stateはサイズをもち、もし関連した情報を持つ新しいコード変換ステートが、コード変換ヒントテーブルに追加される必要があるならば説明する。add_new_transfercoding_stateが「はい」を信号で伝え

るといけないから、以下において新しいコード変換ヒントステートのパラメータのリストが与えられる。パラメータリストのサイズは、1つのコード変換ヒントステートおよびコード変換ヒントステートの数のパラメータの数のため決定される。

【0163】

`remove_transfer<1bit>coding_state`は、もしコード変換ステートが削除できることまたはnotならば、サイズ標示のフラグである。コード変換ステートが削除できるといけないから(整数)以下国の数(サイズ: 削除されるコード変換ステートで与えられる。

【0164】

<1bit>もし図15において定義されるようなより詳細な符号化複雑さ記述体系が使われる必要があるならば、`use_encoding_complexity_description`はサイズとシグナルをもっています。

【0165】

複雑なメタデータの符号化

図15は、コード化複雑さコード変換ヒントメタデータを示し、それは全体の視聴覚コンテンツまたは具体的な視聴覚セグメントに関連できます。以下において、適切な記述子と記述体系が明らかにされる。符号化複雑さメタデータはレートコントロールのために使われて、量子化とビットレート設定を決定する。

【0166】

`use_feature_points`はサイズをもち<1bit>、機能ポイントに基づいた複雑さ評価データの使用を示す。

【0167】

`select_feature_point_method`はタイプをもち<2bits>、機能ポイント方法を選びます。

【0168】

フレームあたり`number_of_new_feature_points`は、図3において示されるようにフレームあたり新しい機能ポイントの数のリストを説明し、サイズである。<nframes*int>。メートル法であるこれは、フレームあたり新しいコンテンツの量

を示する。

【0169】

feature_point_metricsは、1つのセグメント内のフレームあたり新しいfeature pointsに基づく韻律学のリストを説明する。(整数)以下の意味によって、価値の注文されたリストを見せる韻律学：平均、max、min、相違、新しい機能の番号の標準偏差はフレームで示する。

【0170】

use_equation_descriptionは、(整数)フレームあたり符号化複雑さの方程式に基づく記述にポインタである。

【0171】

use_motion_descriptionはタイプをもち<1 bit>、動作基づいた複雑さ記述の使用を示する。

【0172】

select_motion_methodはタイプをもち<4 bits>、動作記述方法を選びます。

【0173】

param_k_motionはタイプをもち<nframes*k*int>、kパラメータをグローバルなパラメータ動作モデルのすべての単一のフレームに記述する。

【0174】

motion_metricsは、韻律学のリストを、動きベクトルのサイズに基づく全体のセグメントに記述する。韻律学が、(整数)価値の注文されたリストであるのが以下の意味によって見せられる。：平均、max、min、バール、大型のブロック動きベクトルのstddev。

【0175】

block_motion_fieldは、am*mブロック寸法決め動作フィールドのすべてのベクトルを説明し、サイズ<n><frames*int*size_x*size_y></>である。<(m*m)>。

【0176】

use_texture_edge_metricsは旗であり、それはセットされる。その時テクスチャまたはエッジ韻律学は使われて、それはサイズをもっています<1 bit>。

【 0 1 7 7 】

`select_texture_edge`はサイズをもち $\langle 4 \text{ bit} \rangle$ 、それは、出現からメートル法であるどのテクスチャーが使われるかを決定する。

【 0 1 7 8 】

`DCT_block_energy`は、1つのブロックのすべてのDCT-係数の合計であり、フレーム内のすべてのブロックのために定義される。それはサイズをもち $\langle \text{size_y} * \text{size_x} * \text{nframes} * \text{int} / 64 \rangle$ 。

【 0 1 7 9 】

`DCT_block_activity_metric`は、1つのブロックけれどもDC係数なしである。すべてのDCT-係数の合計と定義される。それはフレーム内のすべてのブロックのために定義されて、サイズである。 $\langle \text{size_y} * \text{size_x} * \text{nframes} * \text{int} / 64 \rangle$

`DCT_energy_metric`は、韻律学のリストを、個々のブロックの個々のDCT活動力に基づく全体のセグメントに記述する。メートル法が(整数)、価値の注文されたりストであるのが以下の意味によって見せられる。: 平均、max、min、相違、すべての個々のDCTエネルギー韻律学の標準偏差。

【 0 1 8 0 】

記述子のサイズはそうである。 $\langle 6 * \text{int} \rangle$ 。この記述子の代わりのインプリメンテーションは、`DCT_energy_metric`をビデオセグメントのすべての単一のフレームに記述する必要がある。

【 0 1 8 1 】

`DCT_activity_metric`は、韻律学のリストを、個々のブロックの個々のDCT活動に基づく全体のセグメントに記述する。メートル法が(整数)、価値の注文されたりストであるのが以下の意味によって見せられる。: 平均、max、min、相違、すべての個々のDCT活動韻律学の標準偏差。記述子のサイズはそうである。 $\langle 6 * \text{int} \rangle$ 。この記述子の代わりのインプリメンテーションは、`DCT_activity_metric`をビデオセグメントのすべての単一のフレームに記述する必要がある。

【 0 1 8 2 】

ヒントステートメタデータのコード化転送

図 1 6 はコード変換ヒントステートメタデータを示し、それは全体の視聴覚コ

ンテンツまたは具体的な視聴覚セグメントに関連できます。以下において、適切な記述子と記述体系が明らかにされる。

【 0 1 8 3 】

Mは整数タイプであり、I-フレーム/P-フレーム距離を説明する。

【 0 1 8 4 】

~~bitrate_fraction_for_Iはタイプをもち<float>、I-が形成することのために~~
入手可能なこの視聴覚セグメントのために定義されたビットレートの小数部を説明する。

【 0 1 8 5 】

bitrate_fraction_for_Pはタイプをもち<float>、Pフレームのために使われうるこの視聴覚セグメントのために定義されたビットレートの小数部を説明する。B-フレームのためのビットレート小数部は100%にパーセンテージの残りである。

【 0 1 8 6 】

量子化_scale_ratio_I_Pはタイプをもち<float>、間に、量子化スケール（このセグメントのために定義されるような）の関係を示する。I-とP-フレーム。

【 0 1 8 7 】

量子化_scale_ratio_I_Bはタイプをもち<float>、間に、量子化スケール（このセグメントのために定義されるような）の関係を示する。I-とB-フレーム。注意してください。thatビットレート記述子（bitrate_fraction_for_I、bitrate_fraction_for_P）量子化_scale_ratio記述子（量子化_scale_ratio_I_P、量子化_scale_ratio_I_B）または、以下のレートコントロールパラメータは義務的である。

【 0 1 8 8 】

X_I、X_P、X_Bはframe_vbv_complexitiesであり（整数）、それぞれタイプであり、フレームに基づく圧縮目標フォーマット（cfの場合に定義される。図12）。これらおよび以下のVBV（バーチャルなバッファ証明する人）複雑さ調整はオプションで、ソースコンテンツ特徴と目標フォーマット定義に従ってレートコントロール計画を修正するために使用できます。

【0189】

X_I_top、X_P_top、X_B_topは、一番上のフィールドのためのfield_vbv_complexitiesであり(整数)、それぞれタイプであり、フィールドに基づく圧縮目標フォーマット(cfの場合に定義される。図12)。

【0190】

~~X_I_bot、X_P_bot、X_B_botは、~~ボトムフィールドのためのfield_vbv_complexitiesであり(整数)、それぞれタイプであり、フィールドに基づく圧縮目標フォーマット(cfの場合に定義される。図12)。

【0191】

【発明の効果】

本発明によれば、トランスコーディングヒントメタデータの表現に適用可能な方法と装置を実現できる。

【図面の簡単な説明】

【図1】

家庭のネットワークの典型的なコード変換システムにおいて様々な視聴覚ターミナルによってシステム概要を示す図である。

【図2】

コード変換ヒント抽出を示す(グループのピクチャ、GOP)図である。

【図3】

フレームあたり新しい機能ポイントの数に依存しているコード変換ステートの選択のための例を示す図である。

【図4】

3つのステートによってコード変換ヒントの例にステート図を示す図である。

【図5】

転送が、圧縮された、未圧縮であったソースコンテンツからヒントメタデータ抽出をコード化するのを示す図である。

【図6】

ビデオセグメンテーションを示し、コード変換ヒントは選択を述べる図である。

【図7】

新しいビデオセグメントの境界を決定するとすぐに、方法を示す。(または、新しいGOP) 図である。

【図8】

どのようにコード変換ヒントステートを選ぶかにおいてアルゴリズムを示す図である。

【図9】

コード変換ヒントメタデータの構造の組織で概要を与える図である。

【図10】

一般コード変換ヒントメタデータ記述体系の構造の組織を示す図である。

【図11】

ソースフォーマット定義のためのコード変換ヒントメタデータを示す図である。

【図12】

目標フォーマット定義のためのコード変換ヒントメタデータを示す図である。

【図13】

一般コード変換ヒントメタデータ表現を示す図である。

【図14】

セグメントに基づくコード変換ヒントメタデータ表現を示す図である。

【図15】

符号化複雑さコード変換ヒントメタデータを示す図である。

【図16】

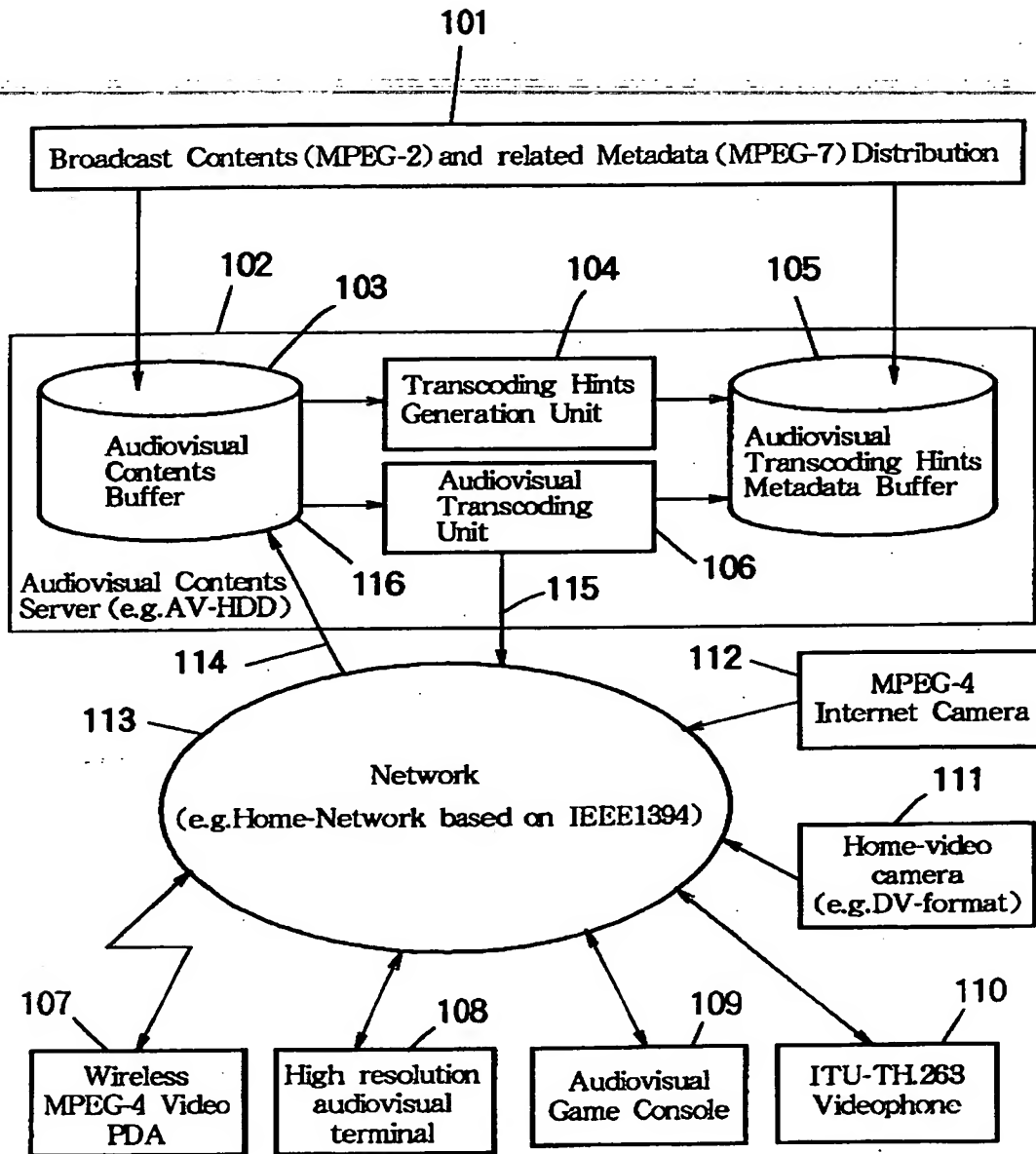
コード変換ヒントステートメタデータを示す図である。

【符号の説明】

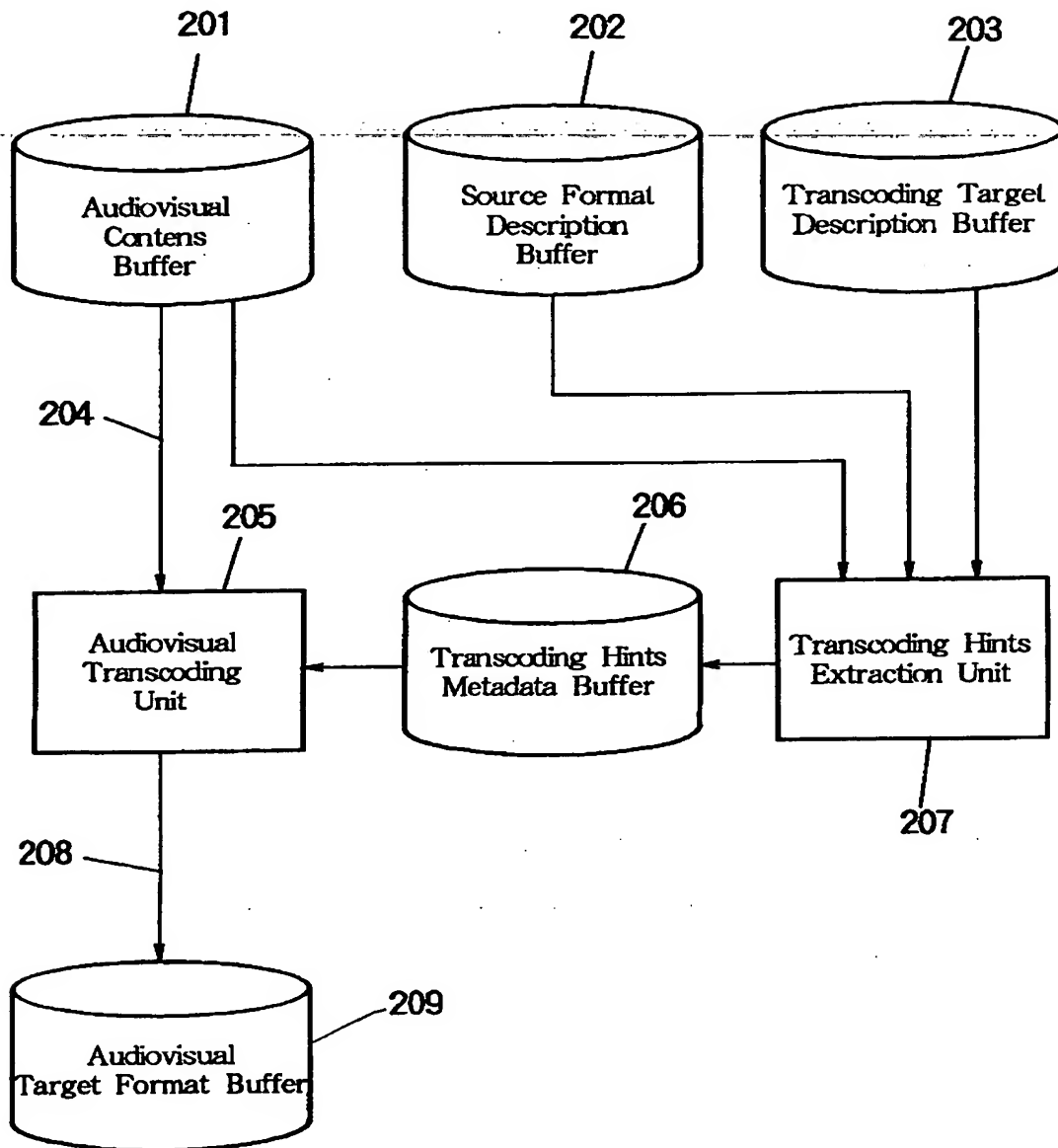
102 ユニット, 103 視聴覚コンテンツ記憶装置, 106 視聴覚
変換ユニット

【書類名】 図面

【図 1】



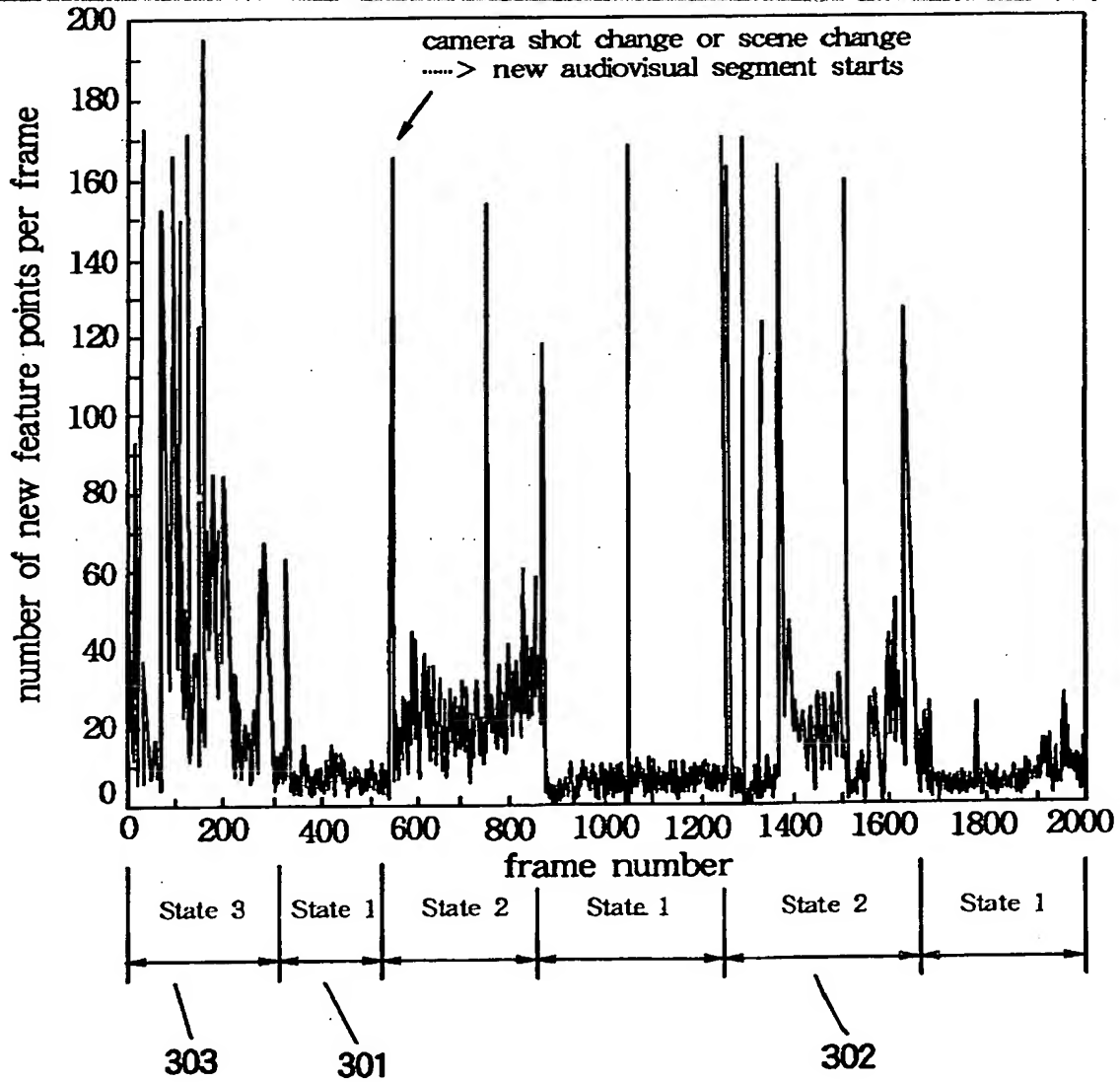
【図 2】



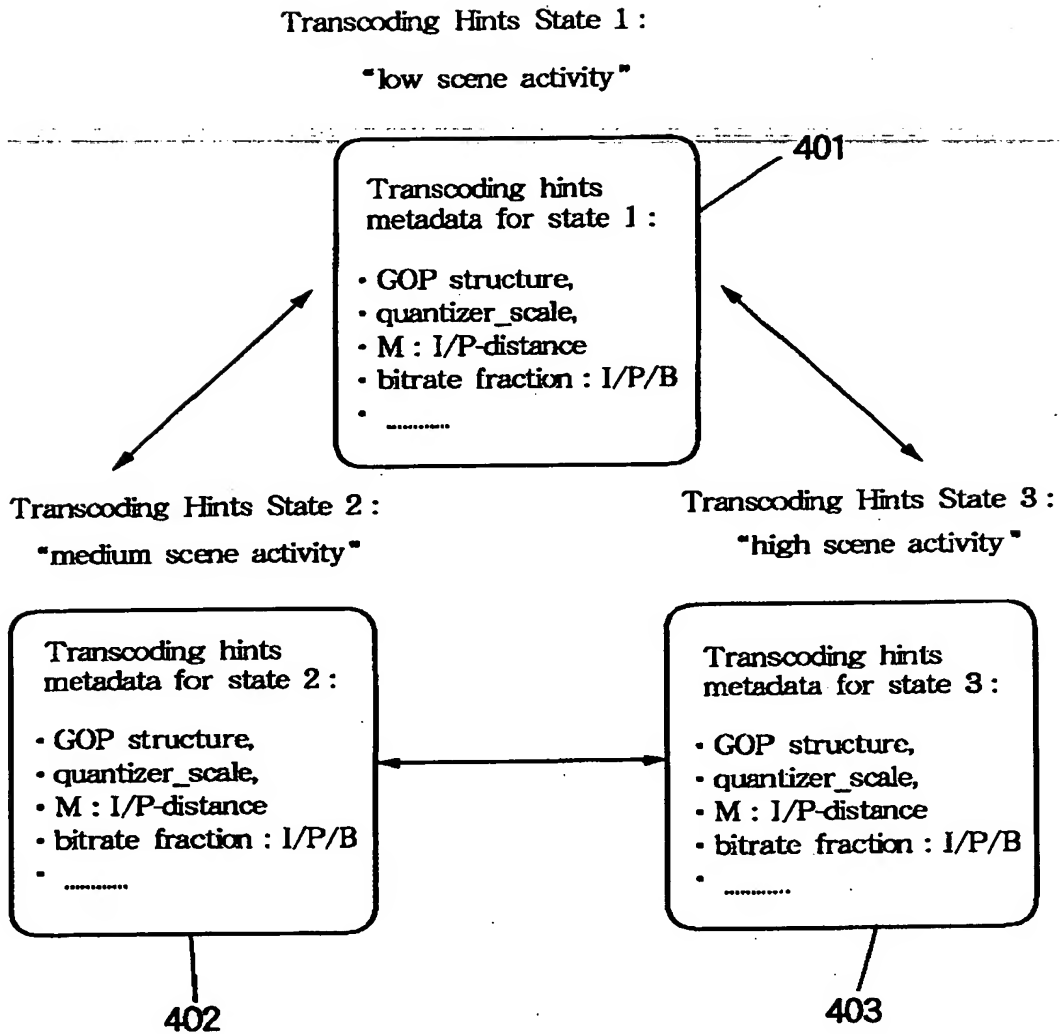
【図 3】

number of new feature pointss per frame

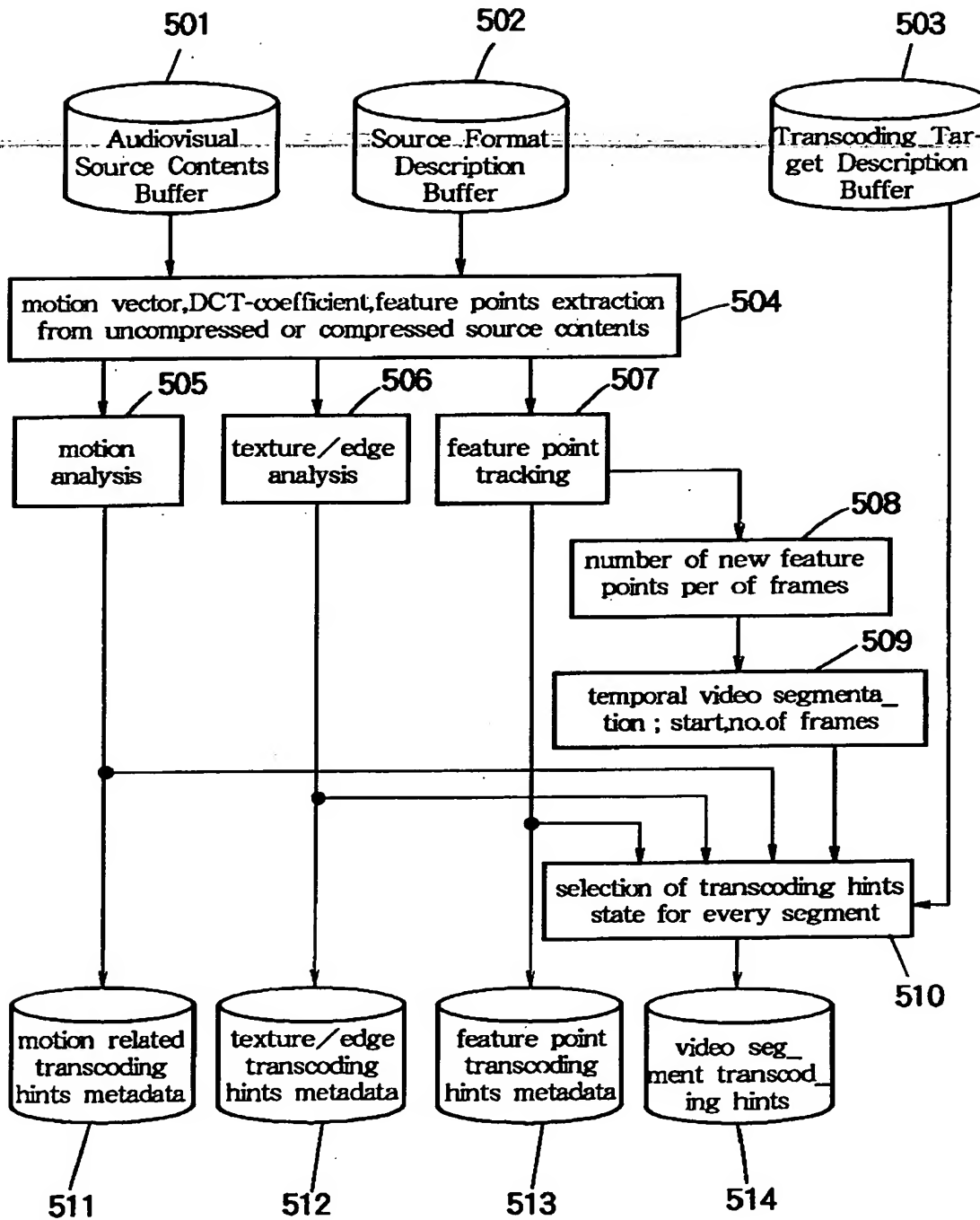
MOEG-7 test sequence : news1.mpg,first 2000 frames



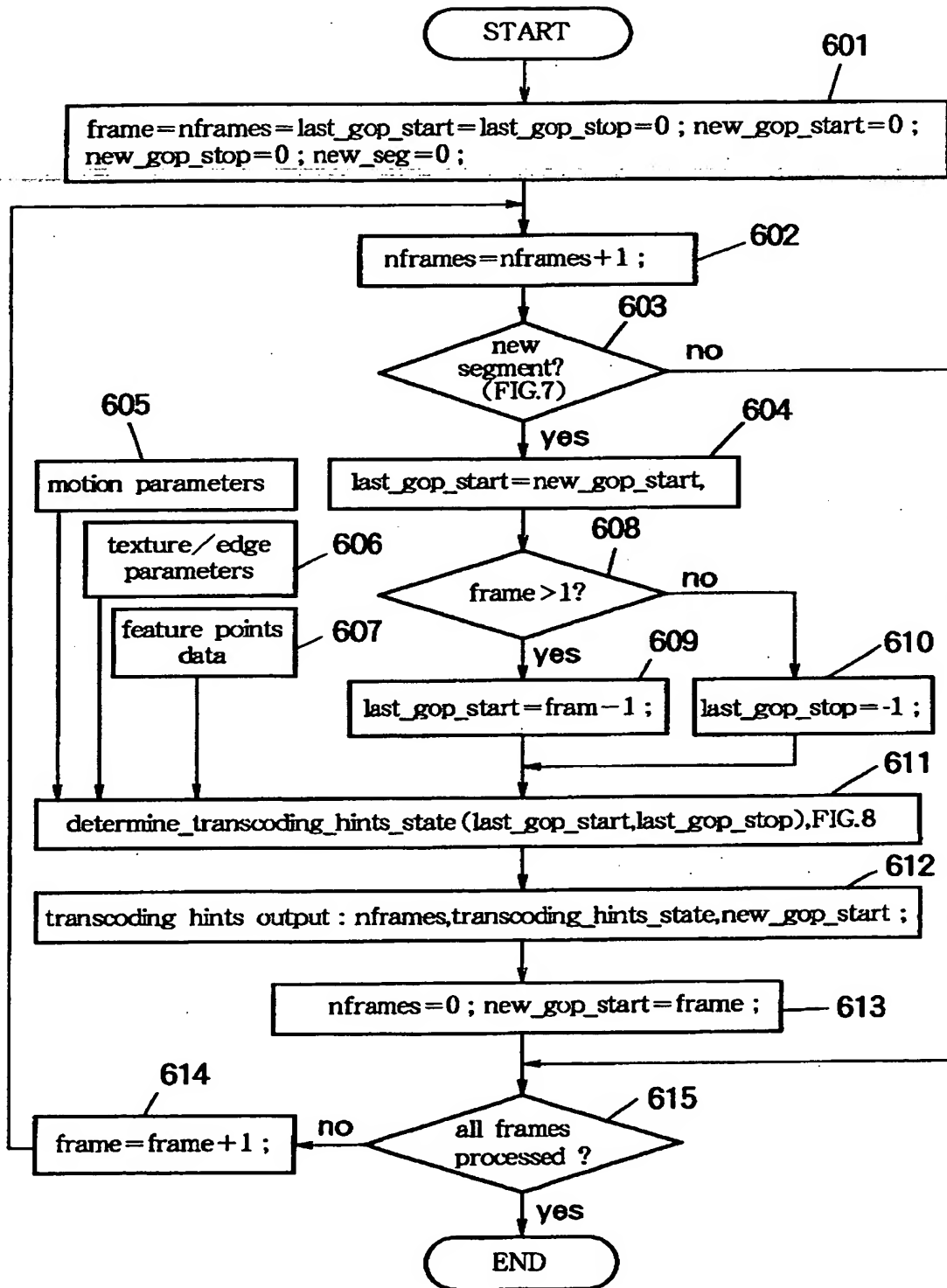
【図 4】



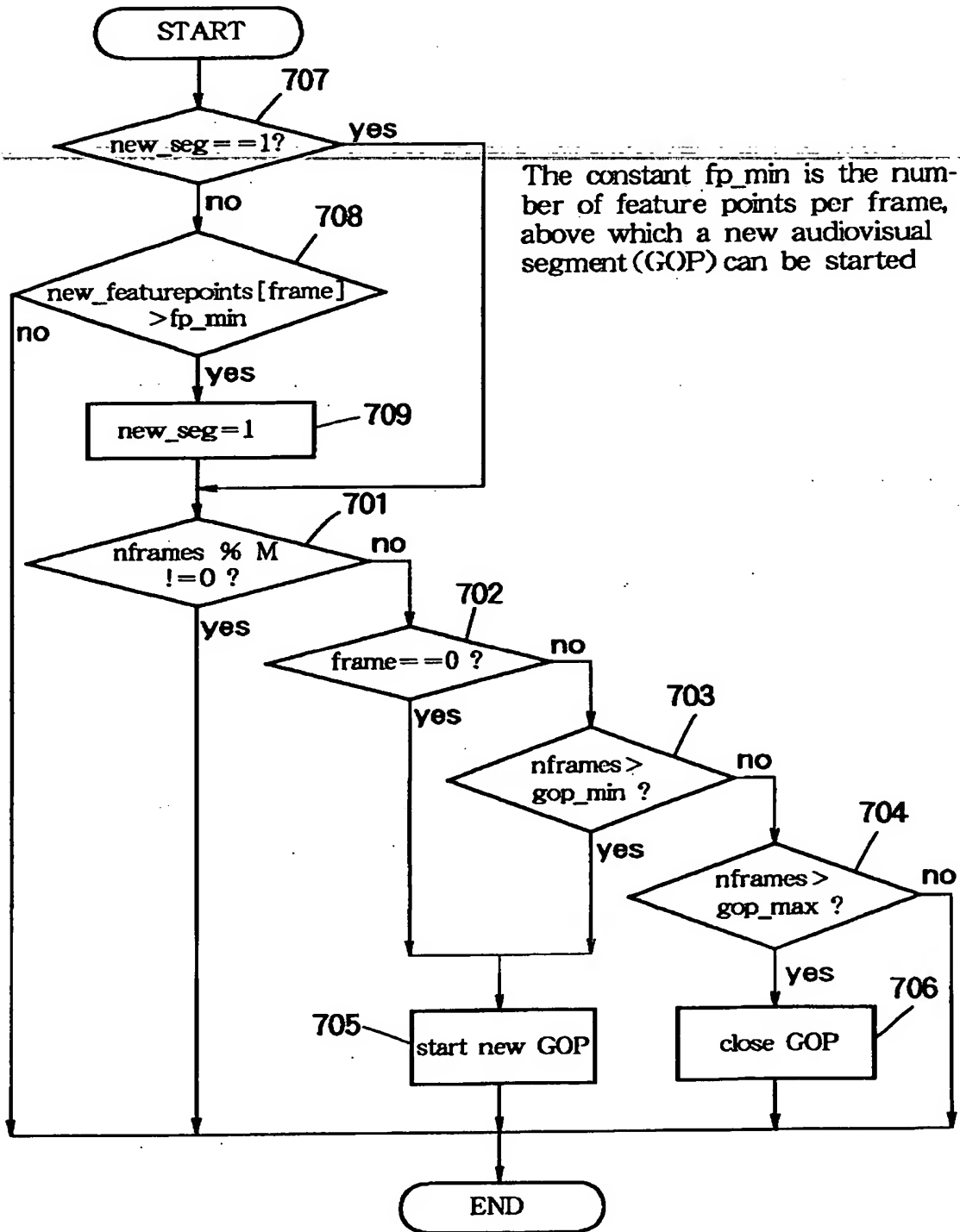
【図5】



【図 6】

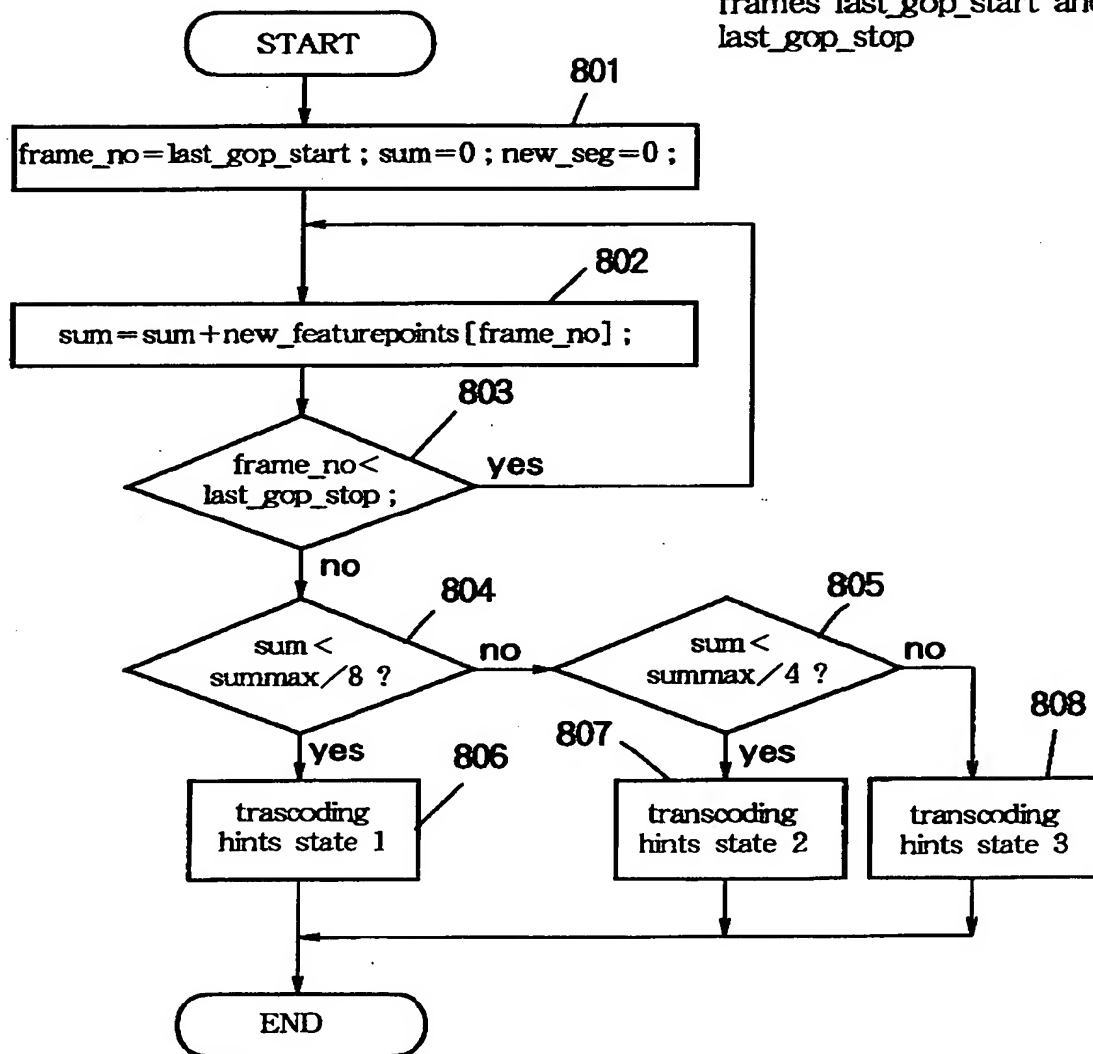


【图 7】

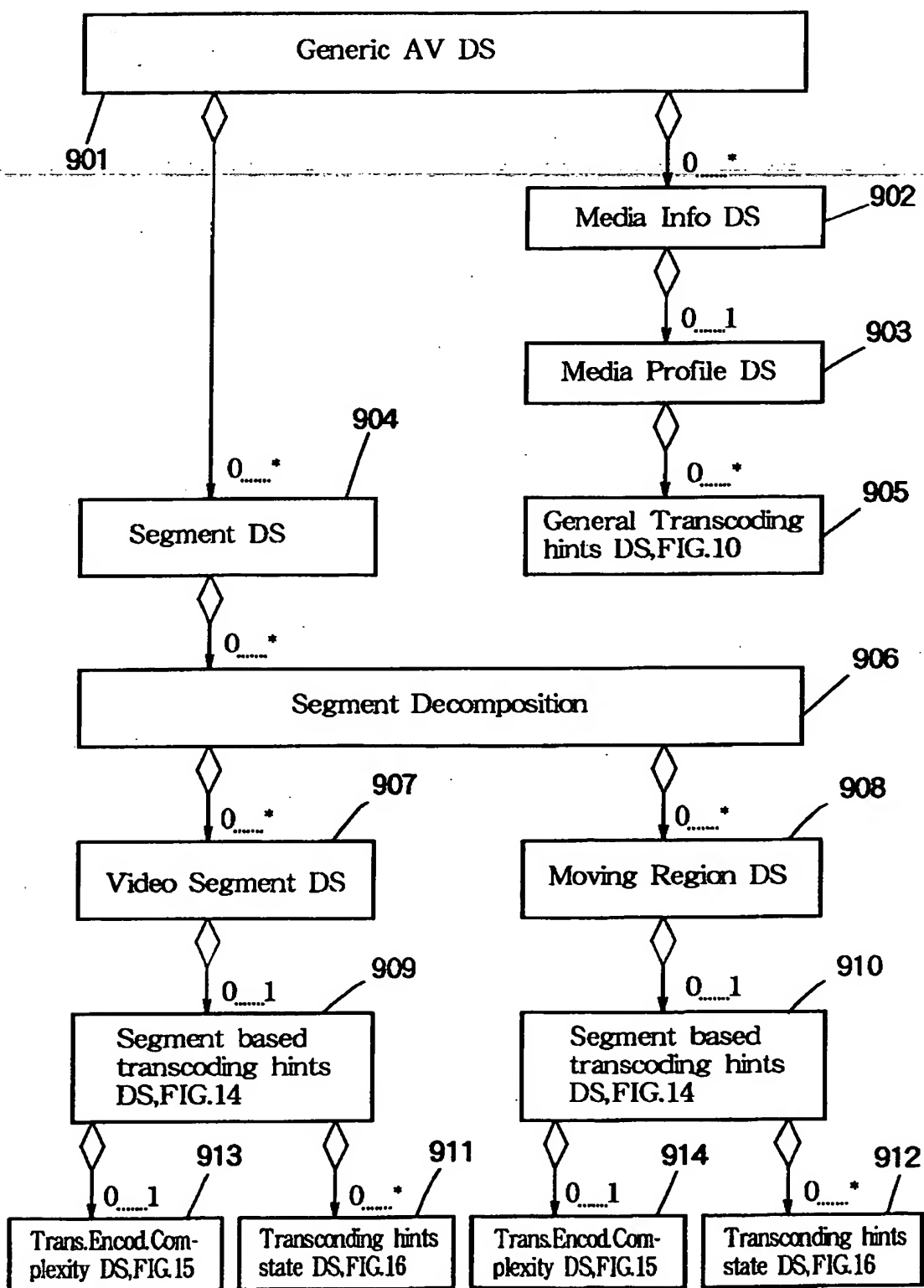


【図 8】

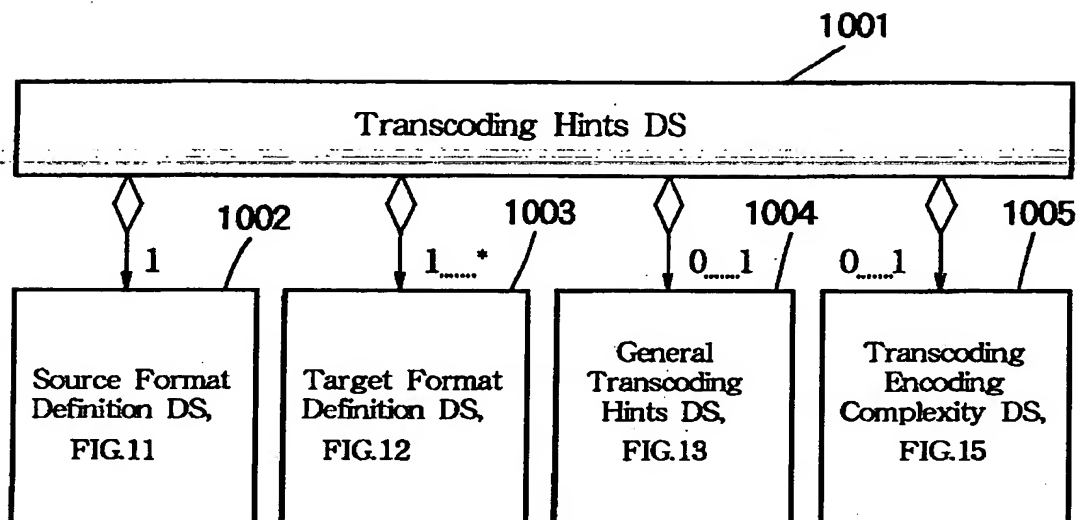
the constant summax
is the maximum
number of featurepoints
which can be tracked
from frame to frame
multiplied by the number
of frames between the
frames last_gop_start and
last_gop_stop



【図 9】



【図10】



【図11】

- 1 : bitrate < int >
- 2 : size_of_pictures < 2*int >
- 3 : number_of_frames_per_second < int >
- 4 : pel_aspect_ratio < float >
- 5 : pel_colour_depth < int >
- 6 : usage_of_progressive_interlaced_format < 1 bit >
- 7 : usage_of_frame_field_pictures < 1bit >
- 8 : compression method < int >
- 9 : one out of list {MPEG-1, MPEG-2, MPEG-4, DV, H.263, H.261,}
- 10 : {further parameters for compression method}
- 11 : GOP_stucture (Runlength coding)

【図 1 2】

```

1 : bitrate < int >
2 : size_of_pictures < 2*int >
3 : number_of_frames_per_second < int >
4 : pel_aspect_ratio < float >
5 : pel_colour_depth < int >
6 : usage_of_progressive_interlaced_format < 1 bit >
7 : usage_of_frame_field_pictures < 1 bit >
8 : compression_method < int >
9 :   one_out_of_list { MPEG-1, MPEG-2, MPEG-4, DV, H.263, H.261, ..... }
10 :   { further parameters for compression method }
11 :   GOP_stucture (Runlength coding)

```

【図 1 3】

```

1 : use_region_of_interest_DS : < 1 bit >
2 :   region of interest DS :
3 :     shape_D : select one of { boundary box D, MB shape D, shape D }
4 :     motion_trajectory_D
5 :
6 : use_editing_effects_transcoding_hints_DS : < 1 bit >
7 :   camera_flash { frame1, frame2, ..... framek } < k*int >
8 :   cross_fading { (start_frame, end_frame), ..... } < k*( < int >, < int > ) >
9 :   black_pictures { (start_frame, end_frame), ..... } < k*( < int >, < int > ) >
10 :   fade_in { (start_frame, end_frame), ..... } < k*( < int >, < int > ) >
11 :   fade_out { (start_frame, end_frame), ..... } < k*( < int >, < int > ) >
12 :   abrupt_change { frame1, frame2, ..... framek } < k*int >
13 :
14 : use_motion_transcoding_hints_DS : < 1 bit >
15 :   number_of_regions : < int >
16 :   for_every_region :
17 :     is_region_rectrangular_shaped (y/n) : < 1 bit >
18 :     if_arbitrarily_shaped : use region D for this region
19 :     describe_parametric_object_motion_for_this_region

```

【図 1 4】

```

1 : start_frame <int>
2 : nframes <int>
3 : l_frame_location :
4 :     select_one_out_of_the_following : <2 bit>
5 :         first_frame (default)
6 :         list of frames {frame1, frame2, ..., frame k} <k*int>
7 :         first_frame_and_every_k_frames <int>
8 :         no_l_frame
9 : quantizer_scale <int>
10 : target_bitrate <int>
11 : target_min_bitrate <int>
12 : target_max_bitrate <int>
13 : use_transcoding_states (y/n) <1 bit>
14 : transcoding_state_nr <int>
15 : add_new_transcoding_state (y/n) <1bit>
16 :     if yes : {list of parameters}
17 : remove_transcoding_state (y/n) <1bit>
18 :     if yes : state_nr <int>
19 : use_encoding_complexity_description (y/n) <1 bit>
20 :     if yes : encoding_complexity_description_scheme

```


【図 1 5】

```

1: use_feature_points {y/n} < 1bit >
2:   select_feature_point_method < 2 bits >
3:     number_of_new_feature_points < nframes * int >
4:     feature_point_metrics {mean, max, min, var, stddev} < 5 * int >
5: use_equation_description {y/n} < 1bit >
6: use_motion_description {y/n} < 1bit >
7: select_motion_method < 4 bits >
8:   param_k_motion < nframes * k * int >
9:   motion_metrics {min, max, sum, var, stddev} < 5 * int >
10:   block_motion_field < nframes * int * size_x_size_y / (m * m) >
11: use_texture_edge_metrics {y/n} < 1bit >
12:   select_texture_edge_metrics < 4 bits >
13:   DCT_block_energy < size_y * size_x * nframes * int / 64 >
14:   DCT_block_activity < size_y * size_x * nframes * int / 64 >
15:   DCT_energy_metric {mean, min, max, sum, var, stddev} < 6 * int >
16:   DCT_activity_metric {mean, min, max, sum, var, stddev} < 6 * int >

```

【図 1 6】

```

1: M: I/P distance <int>
2: bitrate_fraction_for_I <float>
3: bitrate_fraction_for_P <float> /* bitrate_fraction of B is rest to 100% )
4: quantizer_scale_ratio_I_P <float>
5: quantizer_scale_ratio_I_B <float>
6: if_frame: /* see target format transcoding hints */
7:   X_I, X_P, X_B <3*int> /* frame_vbv_complexities */
8: if_field:
9:   X_I_top, X_P_top, X_B_top <3*int> /* field_top_vbv_complexities */
10:  X_I_bot, X_P_bot, X_B_bot <3*int> /* field_bottom_vbv_complexities */

```

【書類名】 要約書

【要約】

【課題】 トランスコーディングヒントメタデータを表すのに好適な方法を提供する。

【解決手段】 コード変換目標ビットストリームのパラメータを記述し、コード変換ヒントメタデータを抽出し、記憶する。セグメントへの視聴覚素材を分離し、視聴覚セグメントに対してコード変換ヒントメタデータを関連させる。

【選択図】 図 1

特2000-068720

出 願 人 履 歴 情 報

識別番号

[000002185]

1. 変更年月日	1990年 8月30日
[変更理由]	新規登録
住 所	東京都品川区北品川6丁目7番35号
氏 名	ソニー株式会社